GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 19:06:22 ; Search time 129.72 Seconds
(without alignments)
6731.630 Million cell updates/sec

Title: US-09-762-311-4_COPY_176_3730
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

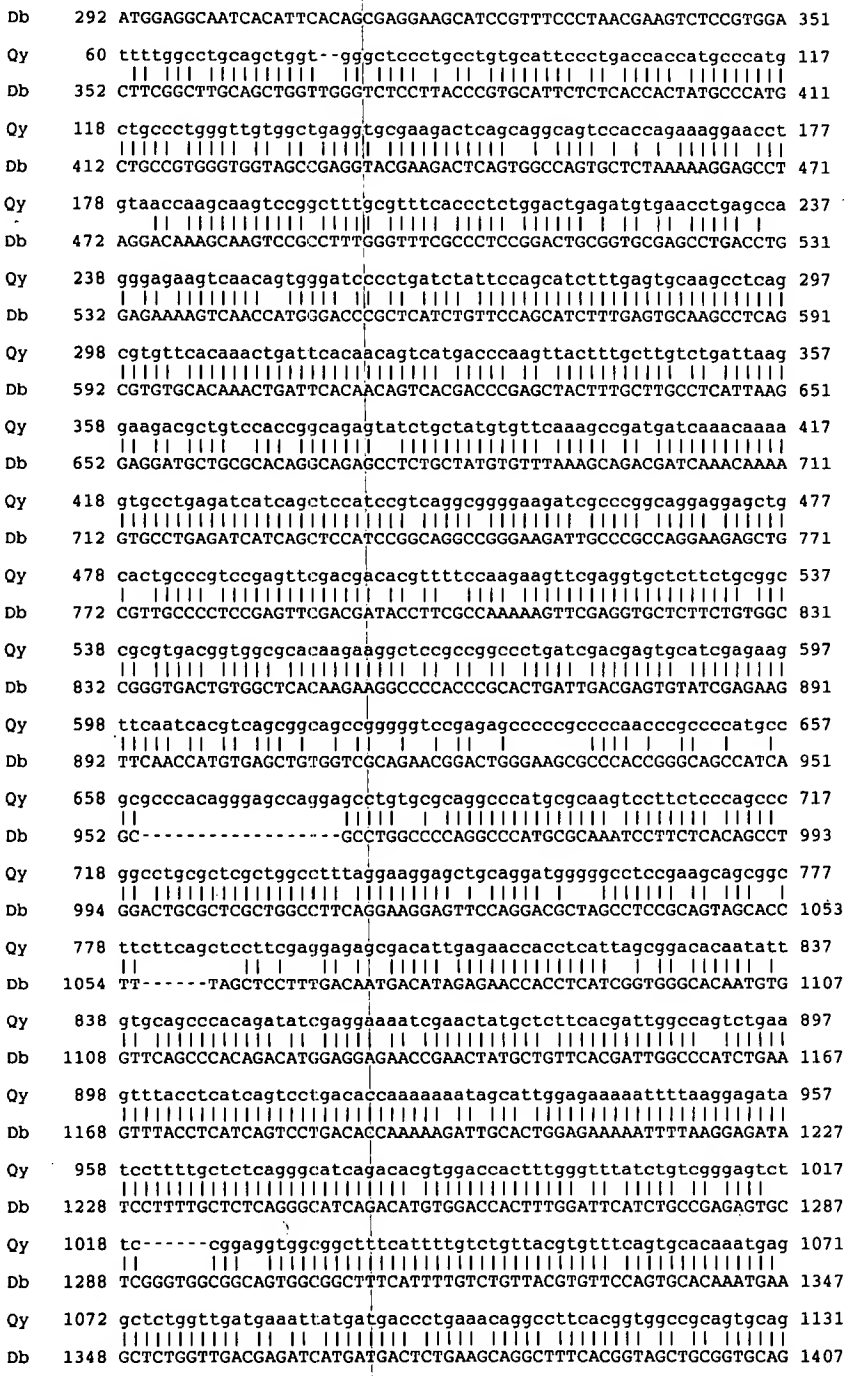
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2510.4	70.6	4039	1	US-08-363-300-1	Sequence 1, Appli
c	2	57.8	1.6	7218	1	US-08-232-463-14	Sequence 14, Appl
	3	45.4	1.3	9515	1	US-08-920-812-13	Sequence 13, Appl
	4	45.4	1.3	9515	1	US-08-920-827-13	Sequence 13, Appl
	5	45.4	1.3	9515	1	US-08-921-177-13	Sequence 13, Appl
	6	45.4	1.3	9515	1	US-08-362-577C-13	Sequence 13, Appl
	7	45.4	1.3	9515	2	US-08-920-828-13	Sequence 13, Appl
	8	45	1.3	8201	1	US-08-253-155A-9	Sequence 9, Appli
	9	41	1.2	1957	1	US-08-295-060-3	Sequence 3, Appli
	10	40.8	1.1	1030	3	US-08-858-003-2	Sequence 2, Appli
	11	40.8	1.1	1030	3	US-09-078-166-2	Sequence 2, Appli
	12	40.8	1.1	1030	4	US-08-997-467-2	Sequence 2, Appli
	13	40.6	1.1	2249	3	US-08-814-052-19	Sequence 19, Appl
	14	40.6	1.1	2279	3	US-08-814-052-17	Sequence 17, Appl
	15	40.6	1.1	2300	3	US-08-814-052-18	Sequence 18, Appl
	16	40.6	1.1	3183	2	US-08-939-218A-1	Sequence 1, Appli
	17	40.6	1.1	3192	1	US-08-940-661A-1	Sequence 1, Appli
	18	40.6	1.1	3192	2	US-09-083-485-1	Sequence 1, Appli
c	19	40.6	1.1	3358	3	US-09-248-571-2	Sequence 2, Appli
	20	40.4	1.1	4134	4	US-09-162-021B-1	Sequence 1, Appli
	21	40	1.1	1962	2	US-08-919-421A-1	Sequence 1, Appli
	22	40	1.1	1962	4	US-09-342-819-1	Sequence 1, Appli
	23	39.2	1.1	3187	5	PCT-US95-06815-1	Sequence 1, Appli
	24	39.2	1.1	3192	1	US-08-706-037-26	Sequence 26, Appl
	25	39.2	1.1	3192	2	US-09-005-397-26	Sequence 26, Appl
	26	39	1.1	2241	2	US-08-838-219B-20	Sequence 20, Appl
	27	39	1.1	2241	3	US-09-233-336A-20	Sequence 20, Appl
	28	39	1.1	2241	3	US-09-233-752A-20	Sequence 20, Appl
	29	39	1.1	2241	4	US-09-402-036-20	Sequence 20, Appl
	30	39	1.1	2370	2	US-08-838-219B-19	Sequence 19, Appl
	31	39	1.1	2370	3	US-09-233-336A-19	Sequence 19, Appl
	32	39	1.1	2370	3	US-09-233-752A-19	Sequence 19, Appl
	33	39	1.1	2370	4	US-09-402-036-19	Sequence 19, Appl
	34	39	1.1	2403	1	US-08-471-033-30	Sequence 30, Appl
	35	39	1.1	2403	2	US-08-471-044-30	Sequence 30, Appl
	36	39	1.1	2403	2	US-08-463-483A-30	Sequence 30, Appl
	37	39	1.1	2403	2	US-08-471-046A-30	Sequence 30, Appl
	38	39	1.1	2403	2	US-08-470-566B-30	Sequence 30, Appl
	39	39	1.1	2403	2	US-08-838-219B-7	Sequence 7, Appli
	40	39	1.1	2403	2	US-08-469-334-30	Sequence 30, Appl
	41	39	1.1	2403	3	US-09-300-529-30	Sequence 30, Appl
	42	39	1.1	2403	3	US-09-233-336A-7	Sequence 7, Appli
	43	39	1.1	2403	3	US-09-233-752A-7	Sequence 7, Appli
	44	39	1.1	2403	4	US-09-402-036-7	Sequence 7, Appli
c	45	38.8	1.1	1069	4	US-09-372-422A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-363-300-1
; Sequence 1, Application US/08363300
; Patent No. 5700927
; GENERAL INFORMATION:
; APPLICANT: Zon, Leonard and Richardson, Paul
; TITLE OF INVENTION: Tbc1 Gene and Uses Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,300
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04590/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4039 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..3829
US-08-363-300-1

Query Match 70.6%; Score 2510.4; DB 1; Length 4039;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 2949; Conservative 0; Mismatches 536; Indels 37; Gaps 8;

Qy 1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59


Db 292 ATGGAGGCAATCACATTCACAGCGAGGAAGCATCCGTTTCCCTAACGAAGTCTCCGTGGA 351
Qy 60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
Db 352 CTTCGGCTTGCAGCTGGTTGGGTCTCCTTACCCGTGCATTCTCTCACCACTATGCCCATG 411
Qy 118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
Db 412 CTGCCGTGGGTGGTAGCCGAGGTACGAAGACTCAGTGGCCAGTGCTCTAAAAAGGAGCCT 471
Qy 178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
Db 472 AGGACAAAGCAAGTCCGCCTTTGGGTTTCGCCCTCCGGACTGCGGTGCGAGCCTGACCTG 531
Qy 238 gggagaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcag 297
Db 532 GAGAAAAGTCAACCATGGGACCCGCTCATCTGTTCCAGCATCTTTGAGTGCAAGCCTCAG 591
Qy 298 cgtgttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
Db 592 CGTGTGCACAAACTGATTCACAACAGTCACGACCCGAGCTACTTTGCTTGCCTCATTAAG 651
Qy 358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
Db 652 GAGGATGCTGCGCACAGGCAGAGCCTCTGCTATGTGTTTAAAGCAGACGATCAAACAAAA 711
Qy 418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
Db 712 GTGCCTGAGATCATCAGCTCCATCCGGCAGGCCGGGAAGATTGCCCGCCAGGAAGAGCTG 771
Qy 478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
Db 772 CGTTGCCCCTCCGAGTTCGACGATACCTTCGCCAAAAAGTTCGAGGTGCTCTTCTGTGGC 831
Qy 538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
Db 832 CGGGTGACTGTGGCTCACAAGAAGGCCCCACCCGCACTGATTGACGAGTGTATCGAGAAG 891
Qy 598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
Db 892 TTCAACCATGTGAGCTGTGGTCGCAGAACGGACTGGGAAGCGCCCACCGGGCAGCCATCA 951
Qy 658 gcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagccc 717
Db 952 GC----------GCCTGGCCCCAGGCCCATGCGCAAATCCTTCTCACAGCCT 993
Qy 718 ggcctgcgctcgctggccttaggaaggagctgcaggatgggggcctccgaagcagcggc 777
Db 994 GGACTGCGCTCGCTGGCCTTCAGGAAGGAGTTCCAGGACGCTAGCCTCCGCAGTAGCACC 1053
Qy 778 ttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatatt 837
Db 1054 TT------TAGCTCCTTTGACAATGACATAGAGAACCACCTCATCGGTGGGCACAATGTG 1107
Qy 838 gtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaa 897
Db 1108 GTTCAGCCCACAGACATGGAGGAGAACCGAACTATGCTGTTCACGATTGGCCCATCTGAA 1167
Qy 898 gtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaattttaaggagata 957
Db 1168 GTTTACCTCATCAGTCCTGACACCAAAAAGATTGCACTGGAGAAAAATTTTAAGGAGATA 1227
Qy 958 tccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtct 1017
Db 1228 TCCTTTTGCTCTCAGGGCATCAGACATGTGGACCACTTTGGATTCATCTGCCGAGAGTGC 1287
Qy 1018 tc------cggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
Db 1288 TCGGGTGGCGGCAGTGGCGGCTTTCATTTTGTCTGTTACGTGTTCCAGTGCACAAATGAA 1347
Qy 1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
Db 1348 GCTCTGGTTGACGAGATCATGATGACTCTGAAGCAGGCTTTCACGGTAGCTGCGGTGCAG 1407
Qy 1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
Db 1408 CAGACGGCTAAGGCACCAGCCCAGCTCTGTGAGGGCTGCCCCTTGCAAGGCCTGCACAAG 1467
Qy 1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
Db 1468 CTCTGCGAAAGGATAGAGGGAATGAATTCATCTAAAACCAAATTAGAACTCCAGAAGCAC 1527
Qy 1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
Db 1528 TTGACCACACTGACCAATCAGGAGCAGGCCACCATATTCGAGGAGGTTCAGAAATTGAGA 1587
Qy 1312 ccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaa 1371
Db 1588 CCAAGAAACGAGCAGCGAGAGAATGAATTAATTATTTCTTTTCTGAGGTGCTTATATGAA 1647
Qy 1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
Db 1648 GAGAAGCAAAAAGAGCACAGCCACACTGGGGCGCCAAAGCAGACACTACAGGTGGCAGCA 1707
Qy 1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
Db 1708 GAGAATATTGGGAGTGACCTGCCACCCAGTGCTAGCCGGTTCAGGTTAGATTCGCTGAAG 1767
Qy 1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1551
Db 1768 AACAGAGCAAAGAGGTCCTTAACAGAGTCCCTAGAGAGCATTCTGTCCCGGGGTAATAAA 1827
Qy 1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
Db 1828 GCCAGAGGCCTGCAGGACCATTCCGCCAGTGTGGATCTGGACAGCTCCACTTCTAGTACT 1887
Qy 1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
Db 1888 CTAAGTAACACCAGCAAAGAGCTGTCCATGGGTGACAAGGAGGCCTTCCCCGTCTCTGAG 1947
Qy 1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
Db 1948 ACCTCCTTCAAGCTCCTTGGCTCCTCAGATGACCTGTCCAGTGACTCAGAGGGCCACATT 2007
Qy 1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
Db 2008 GCAGAAGAGTCTGCCCTGTTGTCACCCCAGCAGGCGTTCAGAAGGAGAGCCAACACCCTG 2067
Qy 1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
Db 2068 AGTCATTTCCCAGTAGAGTGCCCTGCGCCTCCAGAACCTGCCCAGAGCTCTCCAGGGGTC 2127
Qy 1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
Db 2128 TCTCAAAGGAAACTCATGCGGTACCACTCCGTGAGCACAGAGACGCCTCATGAACGCAAG 2187
Qy 1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
Db 2188 GACTTTGAATCCAAAGCAAACCACCTGGGTGACACAGATGGGACCCCCGTGAAGACCCGG 2247
Qy 1972 aggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgat 2031
Db 2248 CGGCACTCGTGGAGACAGCAGATATTCCTTCGAGTGGCCACTCCACAGAAGGCTTGTGAC 2307
Qy 2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
Db 2308 TCCCCGAGCAGATATGAAGATTATTCCGAGCTGGGAGAGCTCCCTCCACGCTCCCCTTTA 2367
Qy 2092 gaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaaaggacatct 2151
Db 2368 GAACCGGTGTGTGAGGACGGCCCATTTGG-CCAGTACAGGAAGAAAAGAGGAAGACGTCA 2426
Qy 2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg-ctgcttagaat 2210
Db 2427 CGCGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTGCAGCAGATCCTGCCTCGTCAGGAT 2486

```
Qy   2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
          |||||||| |||||||||| |||||||||||||| |||||||||||||| ||||| ||
Db   2487 GGAGAAGGAGAATCAGAAGCTACAAGCCTCTGAAAACGATTTGCTGAACAAACGCCTCAA 2546

Qy   2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
          ||| || |||||||||| ||||| |||||||||||||| ||||||||||||||||||||
Db   2547 GCTTGACTATGAAGAAATCACTCCGTGTCTTAAAGAAGTCACTACAGTGTGGGAAAAGAT 2606

Qy   2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
          |||||||||||||||||||| ||||||||||||||||||||||||| ||||||| |||||
Db   2607 GCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAAGTGCACTCAGCTGT 2666

Qy   2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||| ||
Db   2667 TGGGCAAGGTGTGCCACGTCATCACCGAGGTGAGATCTGGAAATTTCTAGCTGAGCAGTT 2726

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
          |||||||||||||   |||||| |||||||||||||||| ||||| |||||||| ||
Db   2727 CCACCTTAAACACCCATTTCCTAGTAAACAGCAGCCAAAGGACGTGCCCTACAAAGAGCT 2786

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaaccttttcc 2570
          | | ||| ||||||| || |||||||| || ||||| || ||||| ||||||||||||||
Db   2787 CCTGAAGAAGCTGACCTCGCAGCAGCACGCCATTCTCATCGACCTCGGGCGAACCTTTCC 2846

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
           ||||| |||||||||||||||||||||||||||||| ||||| || |||||||||||| |
Db   2847 AACACATCCATACTTCTCTGCCCAGCTTGGAGCAGGTCAGCTGTCACTTTACAACATTCT 2906

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
          |||||||||||| ||||| |||||||| || ||||| |||||||||||||||||||| ||
Db   2907 GAAGGCCTACTCGCTTCTGGACCAGGAGGTTGGATACTGCCAAGGTCTCAGCTTTGTGGC 2966

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
          |||||||||||||||||| |||||||||||||||||| || |||||||||| |||||
Db   2967 AGGCATTTTGCTTCTTCACATGAGTGAGGAAGAGGCGTTCAAGATGCTCAAGTTCCTGAT 3026

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3027 GTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTGCAGATCCAGAT 3086

Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          |||||||| ||  || | || ||| |||||||| |||||||||||| |||||||| ||
Db   3087 GTACCAGCTGTCACGGCTCCTCCACGATTACCACCGAGACCTCTACAACCACCTGGAAGA 3146

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          |||||||  ||||||     |||||| || |||||||| |||||| |||| ||||||||
Db   3147 GCACGAGACTGGCCCCCCTACGTACGCGGCTCCCTGGTTTCTCACCGTGTTCGCCTCACA 3206

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          |||||| || || || |||||||||||||||||||||| || ||||||||| |||||||
Db   3207 GTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGATCAGAGGT 3266

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          |||||||||||| |||||||||| ||||| |||||||||||||||||||| ||||||||
Db   3267 CATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTACAGCATGA 3326

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
           ||||| |||||||| || |||||| ||||| | ||| || |||||||| ||| |||| ||
Db   3327 GAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGCCTGGTGCA 3386

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          |||||| |||||||||| |||||| ||||| |||||||||| || ||| | || || ||
Db   3387 GATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTCCAGGCCTA 3446

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||| || |||||||||||| | || || || ||||| || ||||| ||||||||||||||
Db   3447 TGAGGTCGAGTACCACGTCGTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTCAGTGACAA 3506

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          |||||||||||| ||||| |||||||||||||||  || ||||||||||||||||||||
Db   3507 CCAAAGAATGGAGAAATTGGAGAAAACCAACAGCACGTTGCGCAAACAGAACCTTGACCT 3566

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
```

```
          ||| || |||||||||||||||||| |||||||||||||||||| ||||| | |||||
Db   3567 CCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACGGTAGAGAA 3626

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
           || || | ||| ||||| |||||||||||||     ||| ||| | ||  |||||||
Db   3627 ACTTCTTACCAGCGAGAGTAAGCTGAAGCAGCGTGCGCTGACCCTGGAGGTGGAGCG--T 3684

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
          |||||||||||| |||||||||||||||| |||  |||||   |||||||   |  ||
Db   3685 CGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGCACTCCAGA 3744

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
          ||| || ||||| |||| |||||||| ||||| |||| |||
Db   3745 GCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 3786


RESULT   2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-Fls
US-08-232-463-14


  Query Match             1.6%;  Score 57.8;  DB 1;  Length 7218;
  Best Local Similarity   2.9%;  Pred. No. 9.7e-06;
  Matches   11;  Conservative  225;  Mismatches  147;  Indels    0;  Gaps    0;

Qy   1198 gagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacg 1257
          ::::::: :::::::: :::       :::: ::::  ::::  : ::::: :   :: :
Db   1428 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1369
```

```
Qy   1258 acattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgaga 1317
          : :  ::  ::  ::::: ::: ::  :     :::::::   :::::  ::::  ::::
Db   1368 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1309

Qy   1318 aatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaa 1377
          :: ::: :: ::::::: ::: | ::  :          :::: :   : : :::::::::
Db   1308 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1249

Qy   1378 cagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaat 1437
           :::::::: : :   : :  :::::: :::: ::: :  : ::: :: :: ::::::
Db   1248 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1189

Qy   1438 attggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaa 1497
          :  ::::: :::  :  :  :: :  :   ::  :::  ::: : :  :::::: :::
Db   1188 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1129

Qy   1498 gcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccaga 1557
          : :::::::     :: ::::    :::::: :    :    :::: :: ::::  :::
Db   1128 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1069

Qy   1558 ggcctgcaggaacactccatcag 1580
          ::   ||| |  | ||| | | |
Db   1068 RRATCGCAAGCTCCCTCGACCTG 1046


RESULT   3
US-08-920-812-13
; Sequence 13, Application US/08920812
; Patent No. 5763188
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,812
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-920-812-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
          |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
          || |  || | |  | |   ||||   ||||| |  |     || | |||  || ||  |
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
          ||       | |          ||    |      | ||    |      ||     | | |
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagcccccgccccaacccgccccatgccgcgcccacagggagccagga 677
          |     |  ||| |   ||  |   || |   |||  |||  | || | || | |  ||
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
          |     |      | ||  ||    | | ||||||    |   |||||  |  |||||| |
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
            |  || | ||| ||| || |  ||
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   4
US-08-920-827-13
; Sequence 13, Application US/08920827
; Patent No. 5770375
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,827
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-920-827-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
          |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
          || |  || | |  | |   ||||   ||||| |  |     || | |||  || ||  |
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
          ||      |  |          ||     |       | ||     |       ||      | |
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagcccccgcccccaacccgccccatgccgcgcccacagggagccagga 677
          |     |  ||| |   ||  |   || |   |||  |||  | || | || | |  ||
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
          |  |      | ||  ||  |  | | ||||||   |   |||||  |  |||||||  |
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
            |  || | ||| ||| || ||  ||
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   5
US-08-921-177-13
; Sequence 13, Application US/08921177
; Patent No. 5798211
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/921,177
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-921-177-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
          |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
          || |  || | |  | |   ||||   ||||| |  |     || | |||  || ||  |
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
          ||      |  |          ||     |       | ||     |       ||      | |
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagcccccgcccccaacccgccccatgccgcgcccacagggagccagga 677
          |     |  ||| |   ||  |   || |   |||  |||  | || | || | |  ||
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
          |  |      | ||  ||  |  | | ||||||   |   |||||  |  |||||||  |
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
            |  || | ||| ||| || |  ||
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   6
US-08-362-577C-13
; Sequence 13, Application US/08362577C
; Patent No. 5807673
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/362,577C
;      FILING DATE:  27-MAR-1995
```

```
;     CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-362-577C-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
          |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
          || |  || | |  | |   ||||  ||||| |  |     || | |||  || ||  |
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
          ||      |  |          ||    |       | ||     |       ||      | |
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagccccccgccccaacccgccccatgccgcgcccacagggagccagga 677
          |   |  ||| |   ||  |  || |   ||| ||| | || | || | |  ||
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
          |    |      | ||  ||   | | ||||||   |   |||||  |  ||||||| |
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
            |  || | ||| ||| || |  ||
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   7
US-08-920-828-13
; Sequence 13, Application US/08920828
; Patent No. 5853998
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,828
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-920-828-13


  Query Match             1.3%;  Score 45.4;  DB 2;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
          |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
          || |  || | |  | |   ||||  ||||| |  |     || | |||  || ||  |
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
          ||      |  |          ||    |       | ||     |       ||      | |
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagccccccgccccaacccgccccatgccgcgcccacagggagccagga 677
          |   |  ||| |   ||  |  || |   ||| ||| | || | || | |  ||
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
          |    |      | ||  ||   | | ||||||   |   |||||  |  ||||||| |
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
            |  || | ||| ||| || |  ||
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   8
US-08-253-155A-9
; Sequence 9, Application US/08253155A
; Patent No. 5691147
;  GENERAL INFORMATION:
;    APPLICANT:  Gyuris, Jeno
;    APPLICANT:  Draetta, Giulio
;    TITLE OF INVENTION:  CDK4 Binding Proteins
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
```

```
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII(text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/253,155A
;      FILING DATE:  02-JUN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MII-028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8201 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-253-155A-9


  Query Match             1.3%;  Score 45;  DB 1;  Length 8201;
  Best Local Similarity   53.0%;  Pred. No. 0.039;
  Matches   96;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy    2549 ttgaccttgggcgaaccttttcctacacacccatacttctctgcccagcttggagcaggac 2608
           | ||| |  ||   || ' | |  |  ||    | |||    |  |    ||||||    |
Db    2133 TGGACGTGAGGACGACTCTCCGGAACCATGTCTTCTTTAGGGATCGATATGGAGCCAAGC 2192

Qy    2609 agctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatatt 2668
           ||       || | | ||||   |   |||||| ||        ||| ||| ||||| || |
Db    2193 AGAGGGAACTATTCTACATCCTCCTGGCCTATTCGGAGTATAACCCGGAGGTGGGCTACT 2252

Qy    2669 gccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgt 2728
           ||   |  || |||    |  | | | | || || ||| || ||  |||||| || || |
Db    2253 GCAGGGACCTGAGCCACATCACCGCCTTGTTCCTCCTTTATCTGCCTGAGGAGGACGCAT 2312

Qy    2729 t 2729
           |
Db    2313 T 2313


RESULT   9
US-08-295-060-3
; Sequence 3, Application US/08295060
; Patent No. 5659123
;  GENERAL INFORMATION:
;    APPLICANT:  VAN RIE, Jeroen
;    APPLICANT:  JANSENS, Stefan
;    APPLICANT:  PERFEROEN, Marnix
;    TITLE OF INVENTION:  NEW DIABROTICA TOXINS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/295,060
;      FILING DATE:  26-AUG-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  010830-052
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1957 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  13..1947
US-08-295-060-3


  Query Match             1.2%;  Score 41;  DB 1;  Length 1957;
  Best Local Similarity   47.9%;  Pred. No. 0.22;
  Matches  113;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy     488 ccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacgg 547
           || ||  | ||||||       |||   |  ||  |||  ||    | ||  ||    |
Db    1322 CCCAGACCTACGACAGCNNNCGCAACGTGGGCGCTGTGAGCTGGGACAGCATCGACCAGC 1381

Qy     548 tggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacg 607
           || | | | |||    |       | |||||    | |  | || | |  || | ||  |||
Db    1382 TGCCTCCCGAGACCACCGACGAGCCCCTGGAGAAGGGCTACAGCCACCAGCTGAACTACG 1441

Qy     608 tcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcgcccacag 667
           | |   ||  || |  |   |  |||    || ||| |||    |   ||  | ||
Db    1442 TGATGTGCTTCCTGATGCAGGGCAGCAGGGGCACCATCCCAGTGCTGACCTGGACCCACA 1501

Qy     668 ggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctg 723
            ||||  |||     |   || |  |    |||||    || ||||||    ||||
Db    1502 AGAGCGTGGACTTCTTCAACATGATCGACAGCAAGAAGATCACCCAGCTGCCCCTG 1557


RESULT  10
US-08-858-003-2
; Sequence 2, Application US/08858003
; Patent No. 6060234
;  GENERAL INFORMATION:
;    APPLICANT:  Katz, Leonard
;    APPLICANT:  Stassi, Diane L.
;    APPLICANT:  Summers Jr., Richard G.
;    APPLICANT:  Ruan, Xiaoan
;    APPLICANT:  Pereda-Lopez, Ana
;    APPLICANT:  Kakavas, Stephan J.
;    TITLE OF INVENTION:  NOVEL POLYKETIDE DERIVATIVES
;    TITLE OF INVENTION:  AND RECOMBINANT METHODS FOR MAKING SAME
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Rd.
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/858,003
;      FILING DATE:  16-MAY-1979
```

```
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne Casuto
;      REGISTRATION NUMBER:  P-40,943
;      REFERENCE/DOCKET NUMBER:  4952.US.P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (847)-938-3137
;      TELEFAX:  (847)-938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-858-003-2


  Query Match             1.1%;  Score 40.8;  DB 3;  Length 1030;
  Best Local Similarity   50.0%;  Pred. No. 0.18;
  Matches  102;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy    440 tccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcgacg 499
          ||   |||||||    | |||||  | ||      |||  | | | || ||||  | |||
Db    203 TCACACAGGCGGCTCTGTTCGCCGTGGAGACCTCCCTGTTCCGGCTGTTCGAGGCCCACG 262

Qy    500 acacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaaga 559
            |    |  || |  |  ||   |  | || |  ||||   ||||| | ||| |||  |
Db    263 GCCTCGTCCCCGACTACCTCATCGGCCACTCCATCGGCGAAGTGACCGCGGCCCACCTGG 322

Qy    560 aggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcagcc 619
            |     |  |  |||   |||| |||| ||  |     | | | || | ||   |  |
Db    323 CCGGGGTCCTCGATCTGGCGGACGCGTGCGTCCTGGTCGCCCACCGCGGCCGCCTGATGC 382

Qy    620 gggggtccgagagcccccgcccca 643
            | | ||  |  |  | || | ||
Db    383 AGTCGGCCCGGGCCGGCGGCGCGA 406


RESULT  11
US-09-078-166-2
; Sequence 2, Application US/09078166
; Patent No. 6063561
;  GENERAL INFORMATION:
;    APPLICANT:  Katz, Leonard
;    APPLICANT:  Stassi, Diane L.
;    APPLICANT:  Summers Jr., Richard G.
;    APPLICANT:  Ruan, Xiaoan
;    APPLICANT:  Pereda-Lopez, Ana
;    APPLICANT:  Kakavas, Stephan J.
;    TITLE OF INVENTION:  NOVEL POLYKETIDE DERIVATIVES
;    TITLE OF INVENTION:  AND RECOMBINANT METHODS FOR MAKING SAME
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Rd.
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/078,166
;      FILING DATE:  16-MAY-1979
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne Casuto
;      REGISTRATION NUMBER:  P-40,943
;      REFERENCE/DOCKET NUMBER:  4952.US.P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (847)-938-3137
;      TELEFAX:  (847)-938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-078-166-2


  Query Match             1.1%;  Score 40.8;  DB 3;  Length 1030;
  Best Local Similarity   50.0%;  Pred. No. 0.18;
  Matches  102;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy    440 tccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcgacg 499
          ||   |||||||    | |||||  | ||      |||  | | | || ||||  | |||
Db    203 TCACACAGGCGGCTCTGTTCGCCGTGGAGACCTCCCTGTTCCGGCTGTTCGAGGCCCACG 262

Qy    500 acacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaaga 559
            |    |  || |  |  ||   |  | || |  ||||   ||||| | ||| |||  |
Db    263 GCCTCGTCCCCGACTACCTCATCGGCCACTCCATCGGCGAAGTGACCGCGGCCCACCTGG 322

Qy    560 aggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcagcc 619
            |     |  |  |||   |||| |||| ||  |     | | | || | ||   |  |
Db    323 CCGGGGTCCTCGATCTGGCGGACGCGTGCGTCCTGGTCGCCCACCGCGGCCGCCTGATGC 382

Qy    620 gggggtccgagagcccccgcccca 643
            | | ||  |  |  | || | ||
Db    383 AGTCGGCCCGGGCCGGCGGCGCGA 406


RESULT  12
US-08-997-467-2
; Sequence 2, Application US/08997467
; Patent No. 6200813
;  GENERAL INFORMATION:
;    APPLICANT:  Katz, Leonard
;    APPLICANT:  Stassi, Diane L.
;    APPLICANT:  Summers Jr., Richard G.
;    APPLICANT:  Ruan, Xiaoan
;    APPLICANT:  Pereda-Lopez, Ana
;    APPLICANT:  Kakavas, Stephan J.
;    TITLE OF INVENTION:  NOVEL POLYKETIDE DERIVATIVES
;    TITLE OF INVENTION:  AND RECOMBINANT METHODS FOR MAKING SAME
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Rd.
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,467
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/858,003
;      FILING DATE:  16-MAY-1997
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne Casuto
;      REGISTRATION NUMBER:  P-40,943
;      REFERENCE/DOCKET NUMBER:  4952.US.P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (847)-938-3137
;      TELEFAX:  (847)-938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-997-467-2


  Query Match             1.1%;  Score 40.8;  DB 4;  Length 1030;
  Best Local Similarity   50.0%;  Pred. No. 0.18;
  Matches  102;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy    440 tccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcgacg 499
          ||   |||||||    | |||||  | ||      |||  | | | || ||||  | |||
Db    203 TCACACAGGCGGCTCTGTTCGCCGTGGAGACCTCCCTGTTCCGGCTGTTCGAGGCCCACG 262

Qy    500 acacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaaga 559
          |    | || |  |  ||  |  | || |  ||||   ||||| | ||| |||  |
Db    263 GCCTCGTCCCCGACTACCTCATCGGCCACTCCATCGGCGAAGTGACCGCGGCCCACCTGG 322

Qy    560 aggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcagcc 619
            |    |  |  |||   |||| |||| ||  |     | | | || | ||   |  |
Db    323 CCGGGGTCCTCGATCTGGCGGACGCGTGCGTCCTGGTCGCCCACCGCGGCCGCCTGATGC 382

Qy    620 gggggtccgagagcccccgcccca 643
            |  | ||  |  |  | || | |
Db    383 AGTCGGCCCGGGCCGGCGGCGCGA 406


RESULT  13
US-08-814-052-19
; Sequence 19, Application US/08814052
; Patent No. 6015783
;  GENERAL INFORMATION:
;    APPLICANT:  von der Osten, Claus
;    APPLICANT:  Cherry, Joel R.
;    APPLICANT:  Bjornvad, Mads E.
;    APPLICANT:  Vind, Jesper
;    APPLICANT:  Rasmussen, Michael Dolberg
;    TITLE OF INVENTION:  PROCESS FOR REMOVAL OR BLEACHING OF SOILING
;    TITLE OF INVENTION:  OR STAINS FROM CELLULOSIC FABRIC
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6015783o No. 6015783disk of No. 6015783th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,052
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4684.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-814-052-19


  Query Match             1.1%;  Score 40.6;  DB 3;  Length 2249;
  Best Local Similarity   48.9%;  Pred. No. 0.31;
  Matches  109;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy    474 gctgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctg 533
          |||   | ||   |||  |     | || ||    |||       | | |    |  |
Db   1234 GCTCTGCGGCGGCTCCAGGAATCCCTACCCGGCCGCCATCTTCCACTACGCCGGCGCCCC 1293

Qy    534 cggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcga 593
          |||| ||  | |   ||   |    ||||| |||   | ||  | |  |  | | ||
Db   1294 CGGCGGCCCGCCCACGGACGAGGGCAAGGCCCCGGTCGACCACAACTGCCTGGACCTCCC 1353

Qy    594 gaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgcccca 653
            ||  |||| | ||||   | | ||   | ||  ||||  |  | |||| | |||| |
Db   1354 CAACCTCAAGCCCGTCGTGGCCCGCGACGTGCCCCTGAGCGGCTTCGCCAAGCGGCCCGA 1413

Qy    654 tgccgcgcccacagggagccaggagcctgtgcgcaggcccatg 696
            | ||| |   |  | ||  ||  |   | ||| |||| ||
Db   1414 CAACACGCTCGACGTCACCCTCGACACCACGGGCACGCCCCTG 1456


RESULT  14
US-08-814-052-17
; Sequence 17, Application US/08814052
; Patent No. 6015783
;  GENERAL INFORMATION:
;    APPLICANT:  von der Osten, Claus
;    APPLICANT:  Cherry, Joel R.
;    APPLICANT:  Bjornvad, Mads E.
;    APPLICANT:  Vind, Jesper
;    APPLICANT:  Rasmussen, Michael Dolberg
;    TITLE OF INVENTION:  PROCESS FOR REMOVAL OR BLEACHING OF SOILING
;    TITLE OF INVENTION:  OR STAINS FROM CELLULOSIC FABRIC
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6015783o No. 6015783disk of No. 6015783th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,052
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4684.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2279 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-814-052-17


  Query Match             1.1%;  Score 40.6;  DB 3;  Length 2279;
  Best Local Similarity   48.9%;  Pred. No. 0.32;
  Matches  109;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy     474 gctgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctg 533
           |||   | ||   |||  |    | || ||    |||       | | |    |  |
Db    1340 GCTCTGCGGCGGCTCCAGGAATCCCTACCCGGCCGCCATCTTCCACTACGCCGGCGCCCC 1399

Qy     534 cggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcga 593
           ||||  ||  | |   ||   |   |||||  |||   | ||  | |  |  |  |  | ||
Db    1400 CGGCGGCCCGCCCACGGACGAGGGCAAGGCCCCGGTCGACCACAACTGCCTGGACCTCCC 1459

Qy     594 gaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgcccca 653
            ||  |||| | ||||   | | ||  | | ||  ||||  |  | |||| | |||| |
Db    1460 CAACCTCAAGCCCGTCGTGGCCCGCGACGTGCCCCTGAGCGGCTTCGCCAAGCGGCCCGA 1519

Qy     654 tgccgcgcccacagggagccaggagcctgtgcgcaggcccatg 696
             | ||| |   |  | ||  ||  |   | ||| |||| ||
Db    1520 CAACACGCTCGACGTCACCCTCGACACCACGGGCACGCCCCTG 1562


RESULT  15
US-08-814-052-18
; Sequence 18, Application US/08814052
; Patent No. 6015783
;  GENERAL INFORMATION:
;    APPLICANT:  von der Osten, Claus
;    APPLICANT:  Cherry, Joel R.
;    APPLICANT:  Bjornvad, Mads E.
;    APPLICANT:  Vind, Jesper
;    APPLICANT:  Rasmussen, Michael Dolberg
;    TITLE OF INVENTION:  PROCESS FOR REMOVAL OR BLEACHING OF SOILING
;    TITLE OF INVENTION:  OR STAINS FROM CELLULOSIC FABRIC
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6015783o No. 6015783disk of No. 6015783th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,052
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4684.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2300 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-814-052-18


  Query Match             1.1%;  Score 40.6;  DB 3;  Length 2300;
  Best Local Similarity   48.9%;  Pred. No. 0.32;
  Matches  109;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy     474 gctgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctg 533
           |||   | ||   |||  |    | || ||    |||       | | |    |  |
Db    1234 GCTCTGCGGCGGCTCCAGGAATCCCTACCCGGCCGCCATCTTCCACTACGCCGGCGCCCC 1293

Qy     534 cggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcga 593
           ||||  ||  | |   ||   |   |||||  |||   | ||  | |  |  |  |  | ||
Db    1294 CGGCGGCCCGCCCACGGACGAGGGCAAGGCCCCGGTCGACCACAACTGCCTGGACCTCCC 1353

Qy     594 gaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgcccca 653
            ||  |||| | ||||   | | ||  | | ||  ||||  |  | |||| | |||| |
Db    1354 CAACCTCAAGCCCGTCGTGGCCCGCGACGTGCCCCTGAGCGGCTTCGCCAAGCGGCCCGA 1413

Qy     654 tgccgcgcccacagggagccaggagcctgtgcgcaggcccatg 696
             | ||| |   |  | ||  ||  |   | ||| |||| ||
Db    1414 CAACACGCTCGACGTCACCCTCGACACCACGGGCACGCCCCTG 1456


Search completed:  September  9, 2002, 19:07:56
Job time:  16984 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 16:41:06 ; Search time 5191.84 Seconds
(without alignments)
9241.748 Million cell updates/sec

Title: US-09-762-311-4_COPY_176_3730
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	833.6	23.4	1001	10	Z78359	Z78359 HSZ78359 Hu
	2	826.6	23.3	1051	10	BM479578	BM479578 AGENCOURT
	3	819.2	23.0	1007	10	BM016445	BM016445 603641252
	4	782.4	22.0	784	10	BG828092	BG828092 602753463
	5	741.8	20.9	835	10	BG759286	BG759286 602710975
	6	738.8	20.8	800	10	BG746377	BG746377 602703644
	7	737.4	20.7	1121	10	BM476629	BM476629 AGENCOURT
	8	729.8	20.5	1109	10	BM460573	BM460573 AGENCOURT
	9	722.2	20.3	952	10	BE299948	BE299948 600944494
	10	713.8	20.1	838	10	BI907513	BI907513 603065640
	11	705	19.8	807	10	BG120498	BG120498 602346815
	12	701	19.7	832	10	BM014940	BM014940 603640877
	13	689.8	19.4	851	9	AL570425	AL570425 AL570425
	14	683.4	19.2	754	10	BE300005	BE300005 600944594
	15	682.6	19.2	892	10	BE561889	BE561889 601346094
	16	678.6	19.1	686	10	BG769758	BG769758 602744658
	17	675	19.0	1108	10	BG326728	BG326728 602425604
	18	671.8	18.9	978	10	BG029927	BG029927 602297271
	19	669.6	18.8	670	9	AL557623	AL557623 AL557623
c	20	669.4	18.8	857	9	AL544177	AL544177 AL544177
	21	668.4	18.8	750	10	BI084159	BI084159 602869636
	22	664.6	18.7	795	10	BE560184	BE560184 601346896
	23	661	18.6	661	9	AL046787	AL046787 DKFZp434M
	24	661	18.6	948	10	BG258116	BG258116 602379212
	25	659.2	18.5	672	10	BE675153	BE675153 7f03d07.x
	26	652.8	18.4	806	9	AW043925	AW043925 wy82b07.x
	27	648.8	18.3	665	9	AW173375	AW173375 xj78g07.x
	28	647.8	18.2	812	10	BI758489	BI758489 603023334
	29	641	18.0	1161	10	BM466841	BM466841 AGENCOURT
	30	637.8	17.9	773	10	BE561312	BE561312 601344450
c	31	636.4	17.9	862	10	BI910369	BI910369 603068025
	32	635.4	17.9	660	10	BE276812	BE276812 601178843
	33	632.4	17.8	761	10	BF797592	BF797592 602257390
	34	632	17.8	753	10	BF305442	BF305442 601892937
	35	630	17.7	722	10	BE560392	BE560392 601347108
	36	628	17.7	748	10	BG424359	BG424359 602447963
	37	624.4	17.6	627	9	AW772402	AW772402 hn73a08.x
	38	621.2	17.5	626	9	AW239183	AW239183 xb37a10.y
	39	612	17.2	623	10	BE884440	BE884440 601510724
	40	604.4	17.0	649	10	BE513375	BE513375 601315702
	41	597	16.8	605	9	AW732976	AW732976 bb18b11.y
	42	596.2	16.8	622	10	BE514108	BE514108 601315945
	43	592.4	16.7	774	10	BI763670	BI763670 603049558
	44	590.2	16.6	952	10	BG424757	BG424757 602453405
	45	589.4	16.6	758	10	BG166636	BG166636 602339021

ALIGNMENTS

RESULT 1
Z78359/c
LOCUS Z78359 1001 bp mRNA linear EST 28-JUL-1999
DEFINITION HSZ78359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone 3.114 (CEPH) 5', mRNA sequence.
ACCESSION Z78359
VERSION Z78359.1 GI:1495132
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1001)
AUTHORS Neri,C., Albanese,V., Lebre,A.S., Holbert,S., Saada,C., Bougueleret,L., Meier-Ewert,S., LeGall,I., Millasseau,P., Bui,H., Giudicelli,C., Massart,C., Guillou,S., Gervy,P., Poullier,E., Rigault,P., Weissenbach,J., Lennon,G., Chumakov,I., Dausset,J., Lehrach,H., Cohen,D. and Cann,H.M.
TITLE Survey of CAG/CTG repeats in human cDNAs representing new genes: candidates for inherited neurological disorders
JOURNAL Hum. Mol. Genet. 5 (7), 1001-1009 (1996)
MEDLINE 96414310
COMMENT Contact: Neri C.
Fondation Jean Dausset - CEPH
27 Rue Juliette Dodu, 75010 Paris, France
Related sequence: Z78360
5'-sequence (upper strand).
FEATURES Location/Qualifiers
source 1. .1001
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3.114 (CEPH)"
/clone_lib="Human fetal brain S. Meier-Ewert"
/tissue_type="brain"
/dev_stage="fetus"
/note="cDNA library of S. Meier-Ewert, Max Planck Inst.f.Mol.Genetics, Berlin, FRG"
BASE COUNT 237 a 229 c 231 g 272 t 32 others
ORIGIN

```
  Query Match            23.4%;  Score 833.6;  DB 10;  Length 1001;
  Best Local Similarity  93.5%;  Pred. No. 2.2e-197;
  Matches  937;  Conservative    0;  Mismatches   56;  Indels    9;  Gaps    8;

Qy   2146 acatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgctt 2205
          ||||||||||||||||||||||||| | ||||||||||||||||||||| ||||||||
Db   1001 ACATCTCGTGAGCTCCGAGAGCTGTGCCTAAAGGCTATTCTTCAACAGATNCTGCTGCTN 942

Qy   2206 agaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgc 2265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    941 AGAATGGAGAAGGAAAATCAGAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGC 882

Qy   2266 ctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaa 2325
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    881 CTGAAGCTCGATTATGAAGAAATNACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAA 822

Qy   2326 aagatgcttagcactccaggaagat--caaaaattaagtttgacatggaaaaaatgcact 2383
          |||||||||||||||||||||||  || ||||||||||||||||||||||||||||||
Db    821 AAGATGCTTAGCACTCCAGGAAGATCCCANAAATTAAGTTTGACATGGAAAAAATGCACT 762

Qy   2384 cggctgttgggc-aaggtgtgccacgtcatcaccgaggtgaaatctggaaa-tttctagc 2441
          || |||||||||| ||||||||||||||||||||||||||||||||||| ||||||||
Db    761 CGACTGTTGGGCTAAGGTGTGCCACGTCATCACCGAGGTGAAATCTGGAAATTTTCTAGC 702

Qy   2442 tgagcaa-ttccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccat 2500
          ||||||| ||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    701 TGAGCAATTTCCACCGTAAACACCAGTTTCCCAGCANACAGCAGCCANAGGATGTGCCAT 642

Qy   2501 acaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggc 2560
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |
Db    641 ACAAAGAACTCTTNAAGCAGCTGACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGNC 582

Qy   2561 gaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgcttt 2620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    581 GAACCTTTCCTACACACCCATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTT 522

Qy   2621 acaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctca 2680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ACAACATTTTGAAGGCCTACTCACTTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCA 462

Qy   2681 gctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctca 2740
          |||||||| |||||||||||| |||||||||||||||||||||||||| ||| |||||||
Db    461 GCTTTGTACCAGGCATTTTGCGTCTTCATATGAGTGAGGAAGAGGCGTGTAAGATGCTCA 402

Qy   2741 agtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttac 2800
          |||||||||||| |||||||||||  ||  |||| ||||||||||||||||||||||||
Db    401 AGTTTCTGATG-TTGACATGGGCCTNNGGGNACAGNATCGGCCAGACATGATTATTTTAC 343

Qy   2801 agatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatc 2860
          ||| |||||||||||| |||||.|||| | | ||||||||||||||||||||| ||||||
Db    342 AGANCCAGATGTACCANCTCTCGGGGTTNCNTNATGATTACCACAGAGACCTCAACAATC 283

Qy   2861 acctggaggagcacgagatcggccccagcctctacgctg-ccccctggttcctcaccatg 2919
          |||||||||||  |||||||||||||||||||||||| ||||||||||||||||||||||
Db    282 ACCTGGAGGAGGNCGAGATCGGCCCCAGCCTCTACGCTGCCCCCCCTGGTTCCTCACCATG 223

Qy   2920 tttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcag 2979
          |||||||||||||| |||||||||||||  ||||| |||||| ||||||||||||||||
Db    222 TTTGCCTCACAGTTGCCGCTGGGATTCGTATNCAGAGNCTTTGANATGATTTTTCTTCAG 163

Qy   2980 ggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgatt 3039
          |||||||||||||||||||||||||||   ||| ||| |||||| |||||||||  |||||
Db    162 GGAACAGAGGTCATATTTAAAGTGGCGGNAAGNCTGNTGGGAANCCATAAGCCNGTGATT 103

Qy   3040 ctgcagcatgaaaacctagaaa-ccatagttgactttataaaaagcacgctacccaacct 3098
          |||||  |||||||||||||||  ||| ||| ||||||| |||||||||||||||||||
Db    102 CTGCANCATGAAAACCTAGAANGCCAGAGTNGACTTTATGTAAAGCACGCTACCCAACCT 43

Qy   3099 tggcttggtacagatggaa-aagaccatcaatcaggtatttg 3139
          |||| ||||  ||||||||  ||||||||||||||||  ||
Db     42 TGGCGTGGTNNAGATGGAAGNAGACCATCAATCAGGTAGGTG 1
```

```
RESULT   2
BM479578
LOCUS       BM479578                1051 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6464998 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5577183
            5', mRNA sequence.
ACCESSION   BM479578
VERSION     BM479578.1  GI:18528620
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1051)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12330  row: d  column: 16
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .1051
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5577183"
                     /clone_lib="NIH_MGC_92"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      296 a    278 c    234 g    243 t
ORIGIN
```

```
  Query Match            23.3%;  Score 826.6;  DB 10;  Length 1051;
  Best Local Similarity  95.8%;  Pred. No. 1.3e-195;
  Matches  892;  Conservative    0;  Mismatches   34;  Indels    5;  Gaps    4;

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCACCTTAAACACCAGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACT 60

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTAAAGCAGCTGACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCC 120

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TACACACCCATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTT 180

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAAGGCCTACTCACTTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGC 240

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AGGCATTTTGCTTCTTCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGAT 300

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGAT 360
```

```
Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTACCAGCTCTCGAGGTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGA 420

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCACGAGATCGGCCCCAGCCTCTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACA 480

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GTTCCCGCTGGGATTCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGT 540

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CATATTTAAAGTGGCTTTAAGTCTGTTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGA 600

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAACCTAGAAACCATAGTTGACTTTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACA 660

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GATGGAAAAGACCATCAATCAGGTATTTGAAATGGACATCGCTAAACAGTTACAAGCTTA 720

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGAAGTTGAGTACCACGTCCTTCAAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAA 780

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          || |||||||||| |     ||| | |     |||||||||| ||||||||||
Db    781 CC-AAGAATGGATAATTAGAGAAAACCAACAGCAGGCTTACGCAAACGAAACCTTGACCC 839

Qy   3291 ccttgaacagttgcaggt-ggcaaatggtaggatccaaagccttg--aggccaccattga 3347
           | |  |||||||||||| ||| |||||||| |||||||||| ||   | || |||| ||
Db    840 TCCTTGACAGTTGCAGGTGGGCCAATGGTAAGATCCAAAGCCCTGGAGGCCCCCCATGGA 899

Qy   3348 gaagct-cctgagcagtgagagcaagctgaa 3377
          |||||  |||||  |||  ||||  |||||  ||
Db    900 AAAGCTCCCTGAACAGGGAGACCAAGCCTAA 930


RESULT   3
BM016445
LOCUS       BM016445                1007 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603641252F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5417148 5',
            mRNA sequence.
ACCESSION   BM016445
VERSION     BM016445.1  GI:16530799
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1007)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12064  row: h  column: 13
            High quality sequence stop: 822.
FEATURES             Location/Qualifiers
     source          1. .1007
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5417148"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      274 a    246 c    255 g    231 t      1 others
ORIGIN


  Query Match             23.0%;  Score 819.2;  DB 10;  Length 1007;
  Best Local Similarity   95.1%;  Pred. No. 8.9e-194;
  Matches  910;  Conservative    0;  Mismatches   39;  Indels    8;  Gaps    6;

Qy   2532 gcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatacttctctgc 2591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGC 60

Qy   2592 ccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcacttctaga 2651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAGA 120

Qy   2652 ccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttcttcatat 2711
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    121 CCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCANTTTGCTTCTTCATAT 180

Qy   2712 gagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaa 2771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAA 240

Qy   2772 acagtatcggccagacatgattattttacagatccagatgtaccagctctcgaggttgct 2831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCT 300

Qy   2832 tcatgattaccacagagacctctacaatcacctggaggagcacgagatcggccccagcct 2891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCT 360

Qy   2892 ctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagc 2951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGC 420

Qy   2952 cagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtggctttaag 3011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAG 480

Qy   3012 tctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaaccatagttga 3071
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCTGTCGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGA 540

Qy   3072 ctttataaaaagcacgctacccaaccttggcttggtacagatggaaaagaccatcaatca 3131
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTTTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCA 600

Qy   3132 ggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtaccacgtcct 3191
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGTATCTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCT 660

Qy   3192 tcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataa-attag 3250
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    661 TCAAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAACCTTAG 720

Qy   3251 agaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgca-ggtg 3309
          |||||||||||||||| ||||||||||||||||  ||||||||||||||||| |||| |||
Db    721 AGAAAACCAACAGCAG-TTACGCAAACAGAACCTTGGACTCCTTGAACAGTTGCAGGGTG 779

Qy   3310 gcaaatggtaggatccaaagccttgaggccaccattgagaa--gctcctgagcagtgaga 3367
          |||||||||||||||||||||||||||||| |||||||||     || ||||||||||||
Db    780 GCAAATGGTAGGATCCAAAGCCTTGAGGCCCCCATTGAGAAAGCTCCCGGAGCAGTGAGA 839

Qy   3368 gcaag--ctgaagcaggccatgcttacc-ttagaactggagcggtcggccctgctgcaga 3424
```

```
          |||||    |||||||||||||||||| ||||||||||    |   | ||||||||||||
Db    840 GCAAGGCTGGAAGCAGGCCATGCTTACCTTTAGAACTGGGAGCGGTCGGCCTGCTGCAGA 899

Qy   3425 cggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgca 3481
          |||||||||| ||||||  | ||| |    |  ||      |||||||| |||||||
Db    900 CGGTGGAGGAACTGCGGGCGGGGAACCGCAAAGCCCAGCGACGGGAGCCGGAGTGCA 956


RESULT   4
BG828092
LOCUS       BG828092                 784 bp    mRNA    linear   EST 22-MAY-2001
DEFINITION  602753463F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4906191 5',
            mRNA sequence.
ACCESSION   BG828092
VERSION     BG828092.1  GI:14175679
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1808  row: n  column: 16
            High quality sequence stop: 784.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4906191"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      236 a    190 c    182 g    176 t
ORIGIN


  Query Match           22.0%;  Score 782.4;  DB 10;  Length 784;
  Best Local Similarity  99.9%;  Pred. No. 1.3e-184;
  Matches  783;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1859 ggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttg 1918
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGAAACTTATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTG 60

Qy   1919 aatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcatt 1978
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCAAAGCAAACCATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATT 120

Qy   1979 cctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttcca 2038
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTGGAGGCAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCA 180

Qy   2039 gcagatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccag 2098
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    181 GCAGATATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAG 240

Qy   2099 tttgtgaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagc 2158
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTTGTGAAGATGGGCCCTTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGC 300

Qy   2159 tccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaagg 2218
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCCGAGAGCTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGG 360

Qy   2219 aaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgatt 2278
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAATCAGAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATT 420

Qy   2279 atgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagca 2338
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATGAAGAAATTACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCA 480

Qy   2339 ctccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaag 2398
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCCAGGAAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAG 540

Qy   2399 gtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccctta 2458
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGTGCCACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTA 600

Qy   2459 aacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagc 2518
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AACACCAGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGC 660

Qy   2519 agctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacc 2578
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    661 AGCTGACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACAAACC 720

Qy   2579 catacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcct 2638
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCT 780

Qy   2639 actc 2642
           ||||
Db    781 ACTC 784


RESULT   5
BG759286
LOCUS       BG759286                 835 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602710975F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851090 5',
            mRNA sequence.
ACCESSION   BG759286
VERSION     BG759286.1  GI:14069939
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 835)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1693  row: f  column: 19
            High quality sequence stop: 766.
FEATURES             Location/Qualifiers
     source          1. .835
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4851090"
                     /clone_lib="NIH_MGC_48"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      220 a    230 c    214 g    171 t
ORIGIN


  Query Match             20.9%;  Score 741.8;  DB 10;  Length 835;
  Best Local Similarity   98.8%;  Pred. No. 1.9e-174;
  Matches  768;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy   1591 gatagctccctgtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaag 1650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCATCTGTGTGTGAAAAG 60

Qy   1651 gaggccttgcccatctctgagagctcctttaagctcctcggctcctcggaggacctgtcc 1710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCCTCGGAGGACCTGTCC 120

Qy   1711 agtgactcggagagtcatctcccagaagagccagctccgctgtcgccccagcaggccttc 1770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCGCCCCAGCAGGCCTTC 180

Qy   1771 aggaggcgagcaaacaccctgagtcacttccccatcgaatgccaggaacctccacaacct 1830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAGGAACCTCCACAACCT 240

Qy   1831 gcccgggggtccccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcaca 1890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTATCACTCAGTGAGCACA 300

Qy   1891 gagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggt 1950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTTGGTGATTCTGGT 360

Qy   1951 gggactcctgtgaagacccggaggcattcctggaggcagcagatattcctccgagtagcc 2010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATTCCTCCGAGTAGCC 420

Qy   2011 accccgcagaaggcgtgcgattcttccagcagatatgaagattattcagagctgggagag 2070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTCAGAGCTGGGAGAG 480

Qy   2071 cttcccccacgatctccttttagaaccagtttgtgaagatgggccctttggcccccccacca 2130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCCTTTGGCCCCCCCACCA 540

Qy   2131 gaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaa 2190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAAGGCTATTCTTCAA 600

Qy   2191 cagatactgctgcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgat 2250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    601 CAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGCCTCTGAACATGAT 660

Qy   2251 ttgctgaacaagcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagta 2310
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    661 TTGCTGAACAAGCGCCTGAAGCTCGATTATGACGAAATTACTCCCTGTCTTAAAGAAGTA 720

Qy   2311 actacagtgt-gggaaaagatgctta-gcactccaggaagatcaaaaattaagtttg 2365
           |||||||||| ||||| ||||||||| |||||||||||||||||| ||| | ||| |
Db    721 ACTACAGTGTGGGGAACAGATGCTTAGGCACTCCAGGAAGATCAACAATCAGGTTCG 777
```

```
RESULT   6
BG746377
LOCUS       BG746377                 800 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602703644F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4856891 5',
            mRNA sequence.
ACCESSION   BG746377
VERSION     BG746377.1  GI:14057030
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 800)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1708  row: h  column: 12
            High quality sequence stop: 762.
FEATURES             Location/Qualifiers
     source          1. .800
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4856891"
                     /clone_lib="NIH_MGC_15"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      213 a    218 c    203 g    166 t
ORIGIN


  Query Match             20.8%;  Score 738.8;  DB 10;  Length 800;
  Best Local Similarity   98.6%;  Pred. No. 1.1e-173;
  Matches  787;  Conservative    0;  Mismatches    7;  Indels    4;  Gaps    4;

Qy   1502 agagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcc 1561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 AGAGATCTTTAACAGAGTCTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCC 61

Qy   1562 tgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaaca 1621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 TGCAGGAACACTCCATCAGTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACA 121

Qy   1622 ccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttta 1681
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 CCAGCAAAGAGCCATCTGTGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTA 181

Qy   1682 agctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagc 1741
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 AGCTCCTCGGCTCCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGC 241

Qy   1742 cagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcc 1801
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 CAGCTCCGCTGTCGCCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCC 301
```

```
Qy   1802 ccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagga 1861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 CCATCGAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGA 361

Qy   1862 aacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaat 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 AACTTATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAAT 421

Qy   1922 ccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcct 1981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 CCAAAGCAAACCATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCT 481

Qy   1982 ggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagca 2041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 GGAGGCAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCA 541

Qy   2042 g-atatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2100
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 GCATATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTT 601

Qy   2101 tgtgaa-gatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagct 2159
          |||||| ||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    602 TGTGAACGATGGGCCCTTTGGCCCCCCACCAGAGGAACAGAAAAGGACATCTCGTGAGCT 661

Qy   2160 ccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatgga-gaagg 2218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    662 CCGAGAGCTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGACGAAGG 721

Qy   2219 aaaatcagaagctccaagcctctgaaaatgatttgctg-aacaagcgcctgaagctcgat 2277
          ||||||||||||||||| |||||| |||||||||| ||| ||||||||||||||||||||
Db    722 CAAATCAGAAGCTCCAAGCTTCTGAACATGATTTGCTGAAACCAGCGCCTGAAGCTCGAT 781

Qy   2278 tatgaagaaattactccc 2295
          |||||||||| ||||||
Db    782 CATGAAGAAATCACTCCC 799
```

RESULT 7
BM476629
LOCUS BM476629 1121 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6476337 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5553589 5', mRNA sequence.
ACCESSION BM476629
VERSION BM476629.1 GI:18525671
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1121)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12270 row: m column: 14
High quality sequence start: 6
High quality sequence stop: 636.
FEATURES Location/Qualifiers
source 1. .1121
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5553589"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 312 a 277 c 280 g 251 t 1 others
ORIGIN

Query Match 20.7%; Score 737.4; DB 10; Length 1121;
Best Local Similarity 96.1%; Pred. No. 2.8e-173;
Matches 799; Conservative 0; Mismatches 26; Indels 6; Gaps 4;

```
Qy    830 acaatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggcc 889
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     31 ACTATATTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGATTGGCC 90

Qy    890 agtctgaagtttacctcatcagtcctgacaccaaaaaaaatagcat-tggagaaaaatttt 948
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     91 AGTCTGAAGTTTACCTCATCAGTCCTGACACCAAAAAAATAGCATNTGGAGAAAAATTTT 150

Qy    949 aaggagatatccttttgctctcagggcatcagacacgtggaccactttgggtttatctgt 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151 AAGGAGATATCCTTTTGCTCTCAGGGCATCAGACACGTGGACCACTTTGGGTTTATCTGT 210

Qy   1009 cgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaat 1068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211 CGGGAGTCTTCCGGAGGTGGCGGCTTTCATTTTGTCTGTTACGTGTTTCAGTGCACAAAT 270

Qy   1069 gaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtg 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271 GAGGCTCTGGTTGATGAAATTATGATGACCCTGAAACAGGCCTTCACGGTGGCCGCAGTG 330

Qy   1129 cagcagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcac 1188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    331 CAGCAGACAGCTAAGGCGCCAGCCCAGCTGTGTGAGGGCTGCCCCCTGCAAAGCCTGCAC 390

Qy   1189 aagctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaag 1248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 AAGCTCTGTGAGAGGATAGAGGGAATGAATTCTTCCAAAACAAAACTAGAACTGCAAAAG 450

Qy   1249 cacctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattg 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 CACCTGACGACATTAACCAATCAGGAGCAGGCGACTATTTTTGAAGAGGTTCAGAAATTG 510

Qy   1309 agaccgagaaatgagcagcgagagaatgaattgattatttctttttctgagatgtttatat 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 AGACCGAGAAATGAGCAGCGAGAGAATGAATTGATTATTTCTTTTCTGAGATGTTTATAT 570

Qy   1369 gaagagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggca 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 GAAGAGAAACAGAAAGAACACATCCATATTGGGGAGATGAAGCAGACATCGCAGATGGCA 630

Qy   1429 gcagagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctg 1488
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    631 GCAGAGAATATTGGAAGTGAATTACCACCCAGTGCCCCTCGATTTAGGCTAGATATGCTG 690

Qy   1489 aaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaat 1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    691 AAAAACAAAGCAAAGAGATCTTTAACAGAGTCTTTAGAAAGTATTTTGTCCCGGGGTAAT 750

Qy   1549 aaagccagaggcctgcaggaacactccatcagtgt---ggatctggatagctcccct-gtc 1604
          |||||||||||||||||| |  |  |      ||   ||||||||||||||||||| |||
Db    751 AAAGCCAGAGGCCTGCAGGGAACCCTCCATCCCGTGTGGGATCTGGATAGCTCCCTGGTC 810

Qy   1605 tagtaca-ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggagg 1654
          ||||||| |||||||||||||||||     |     |||||| |||   ||
Db    811 TAGTACATTTAAGTAACACCAGCAAAAGAGCCATCCTGTGTGTAAAAAGGG 861
```

RESULT 8
BM460573
LOCUS BM460573 1109 bp mRNA linear EST 05-FEB-2002

```
DEFINITION  AGENCOURT_6421241 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5531908
            5', mRNA sequence.
ACCESSION   BM460573
VERSION     BM460573.1  GI:18509613
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1109)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12214  row: f  column: 05
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .1109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5531908"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT      301 a    311 c    258 g    239 t
ORIGIN
```

```
  Query Match             20.5%;  Score 729.8;  DB 10;  Length 1109;
  Best Local Similarity   97.5%;  Pred. No. 2.2e-171;
  Matches  773;  Conservative    0;  Mismatches   17;  Indels    3;  Gaps    3;

Qy   1340 tgattatttcttttctgagatgtttatatgaagagaaacagaaagaacacatccatattg 1399
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     13 TGATTATTTCTTTTCTGAGATGTTTATATGAAGAGAAACAGAAAGAACACATCCATATTG 72

Qy   1400 gggagatgaagcagacatcgcagatggcagcagagaatattggaagtgaattaccaccca 1459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     73 GGGAGATGAAGCAGACATCGCAGATGGCAGCAGAGAATATTGGAAGTGAATTACCACCCA 132

Qy   1460 gtgccactcgatttaggctagatatgctgaaaaacaaagcaaagagatctttaacagagt 1519
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 GTGCCACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAGAGATCTTTAACAGAGT 192

Qy   1520 ctttagaaagtattttgtcccggggtaataaagccagaggcctgcaggaacactccatca 1579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 CTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTGCAGGAACACTCCATCA 252

Qy   1580 gtgtggatctggatagctccctgtctagtacattaagtaacaccagcaaagagccatctg 1639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 GTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCATCTG 312

Qy   1640 tgtgtgaaaaggaggccttgcccatctctgagagctcctttaagctcctcggctcctcgg 1699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 TGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCCTCGG 372

Qy   1700 aggacctgtccagtgactcggagagtcatctcccagaagagccagctccgctgtcgcccc 1759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 AGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCGCCCC 432

Qy   1760 agcaggccttcaggaggcgagcaaacaccctgagtcacttccccatcgaatgccaggaac 1819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    433 AGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAGGAAC 492

Qy   1820 ctccacaacctgcccgggggtccccgggggtttcgcaaaggaaacttatgaggtatcact 1879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    493 CTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTATCACT 552

Qy   1880 cagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatcttg 1939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    553 CAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTTG 612

Qy   1940 gtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatattcc 1999
          |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db    613 GTGATTCTGGTGGGACTCCTGTGAAGACCCCGGAGCATTCCTGGAGGCAGCAGATATTCC 672

Qy   2000 tccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattcag 2059
          ||||||||||||||||||||||||||||| ||||||||||||||||||||| |||||||| |
Db    673 TCCGAGTAGCCACCCCGCAGAAGGCGGGCGATTCTTCCAGCAGATATGGAGATTATTCCG 732

Qy   2060 agct-gggagagcttcccccacgatc-tcctttagaaccagtttgt-gaagatgggccct 2116
          |||| |||||||||||||| | | | ||| |||||||||||||| ||||||||| ||
Db    733 AGCTGGGGAGAGCTTCCCCCCCCAACTTCCCTTAGAACCAGTTTGTGGAAGATGGGGCCC 792

Qy   2117 ttggccccccacc 2129
          || |  | || ||
Db    793 TTTGGGCGCCTCC 805
```

```
RESULT   9
BE299948
LOCUS       BE299948                952 bp    mRNA    linear   EST 20-JUL-2000
DEFINITION  600944494F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960782 5',
            mRNA sequence.
ACCESSION   BE299948
VERSION     BE299948.1  GI:9183696
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 952)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM54  row: k  column: 23
            High quality sequence stop: 709.
FEATURES             Location/Qualifiers
     source          1. .952
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2960782"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      249 a    240 c    250 g    212 t      1 others
ORIGIN
```

```
  Query Match            20.3%;  Score 722.2;  DB 10;  Length 952;
  Best Local Similarity  97.2%;  Pred. No. 1.7e-169;
  Matches  766;  Conservative    0;  Mismatches   19;  Indels    3;  Gaps    3;

Qy   1879 tcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatctt 1938
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTT 60

Qy   1939 ggtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatattc 1998
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATTC 120

Qy   1999 ctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattca 2058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTCA 180

Qy   2059 gagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgggccctttt 2118
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCCTTT 240

Qy   2119 ggcccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaag 2178
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAAG 300

Qy   2179 gctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaagcc 2238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGCC 360

Qy   2239 tctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccctgt 2298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCCTGT 420

Qy   2299 cttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaaatt 2358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAAATT 480

Qy   2359 aagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcaccga 2418
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    481 AAGTTTGACATGGAAAAAATGCACTCGGCTG-TGGGCAAGGTGTGCCACGTCATCACCGA 539

Qy   2419 ggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcccagcaaa 2478
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 GGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGCAAA 599

Qy   2479 cagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcagcat 2538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAGCAT 659

Qy   2539 gcgattctta-ttgaccttgggcgaaccttcctacacacccatacttctctgcccagct 2597
          |||||||||| |||||| ||||||||||||||||||||||||||||||||| | |||||
Db    660 GCGATTCTTATTTGACCTGNGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAG-T 718

Qy   2598 tggagcaggacagctatcgctttacaacattttgaaggcctactcacttctagaccagga 2657
          |||||||||||||||||||||||||||||||   ||| ||||||   |      |||||
Db    719 TGGAGCAGGACAGCTATCGCTTTACAACATTTGTGAGGGCTACTCCACTTCTAGACAGGA 778

Qy   2658 agtgggat 2665
             || ||
Db    779 GTGGGTAT 786


RESULT  10
BI907513
LOCUS       BI907513                838 bp    mRNA    linear   EST 16-OCT-2001
DEFINITION  603065640F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214945 5',
            mRNA sequence.
ACCESSION   BI907513
VERSION     BI907513.1  GI:16170347
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11539  row: o  column: 10
            High quality sequence start: 4
            High quality sequence stop: 780.
FEATURES             Location/Qualifiers
     source          1. .838
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5214945"
                     /clone_lib="NIH_MGC_118"
                     /tissue_type="leukocyte"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated  adult donors.  Library is oligo-dT primed
                     and directionally cloned  (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb.  Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
BASE COUNT      237 a    199 c    191 g    211 t
ORIGIN


  Query Match            20.1%;  Score 713.8;  DB 10;  Length 838;
  Best Local Similarity  97.0%;  Pred. No. 2e-167;
  Matches  812;  Conservative    0;  Mismatches   17;  Indels    8;  Gaps    8;

Qy   2226 gaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaaga 2285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 GAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGA 68

Qy   2286 aattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactccagg 2345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 AATTACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGG 128

Qy   2346 aagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgcc 2405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 AAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCC 188

Qy   2406 acgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacacca 2465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 ACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCA 248

Qy   2466 gtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctgac 2525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 GTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGAC 308

Qy   2526 ttcccagcagcatgcgattcttattgaccttgggcgaaccttcctacacacccatactt 2585
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    309 TTCCCAGCAGCATGCGATTCTTA-TGACCTTGGGCGAACCTTTCCTACACACCCATACTT 367

Qy   2586 ctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcact 2645
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    368 CTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACA-TTTGAAGGCCTACTCACT 426

Qy   2646 tctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttct 2705
          |||||||||||||||||||| |||||||||||||||| |||||||||| |||||||||
Db    427 TCTAGACCAGGAAGTGGGATA-TGCCAAGGTCTCAGC-TTGTAGCAGGCA-TTTGCTTCT 483
```

```
Qy   2706 tcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggct 2765
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    484 TCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATG-TTGACATGGGGCT 542

Qy   2766 gcggaaacagtatcggccagacatgattatttacagatccagatgtaccagctctcgag 2825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 GCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGAG 602

Qy   2826 gttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggccc 2885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 GTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCC 662

Qy   2886 cagcctctacgctg-ccccctggttcctcaccatgtttgcctcacagttcccgctgggat 2944
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    663 CAGCCTCTACGCTGCCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGAT 722

Qy   2945 tcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtgg 3004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    723 TCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGG 782

Qy   3005 ctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaa 3061
           ||||| ||   ||||||| |||||||||   |         ||||||||||||||
Db    783 -TTTAAATCCTGTGGGAAGGCATAAGCCCTGATTCTGCAGCATTGAAAACCTAGAAA 838


RESULT 11
BG120498
LOCUS       BG120498                807 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602346815F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4441923 5',
            mRNA sequence.
ACCESSION   BG120498
VERSION     BG120498.1  GI:12614007
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10213  row: n  column: 04
            High quality sequence stop: 682.
FEATURES             Location/Qualifiers
     source          1. .807
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4441923"
                     /clone_lib="NIH_MGC_90"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      231 a    197 c    189 g    190 t
ORIGIN


  Query Match             19.8%;  Score 705;  DB 10;  Length 807;
  Best Local Similarity   96.1%;  Pred. No. 3.1e-165;
  Matches  766;  Conservative    0;  Mismatches   25;  Indels    6;  Gaps    4;
```

```
Qy   1992 gatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaaga 2051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGA 60

Qy   2052 ttattcagagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgg 2111
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGG 120

Qy   2112 gccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtg 2171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCTTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTG 180

Qy   2172 gcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagct 2231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCT 240

Qy   2232 ccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattac 2291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTAC 300

Qy   2292 tccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatc 2351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATC 360

Qy   2352 aaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtca 2411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCA 420

Qy   2412 tcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcc 2471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCC 480

Qy   2472 cagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttccca 2531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCA 540

Qy   2532 gcagcatgcgattcttattgaccttgggcgaacctttcctacaca-cccatacttctctg 2590
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    541 GCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACAGCCCATACTTCTCTG 600

Qy   2591 cccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcacttctag 2650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAG 660

Qy   2651 accaggaagtgggatattgcc--aaggtctcagcttt-gtagcaggcattttgcttcttc 2707
          |||||||||||||||||||  |||   |||||||| || |||||| | ||||||| ||
Db    661 ACCAGGAAGTGGGATATTGCCCAAAGTCTTCAGCTTTGGTGGCAGGCCTCTTGCTTCCTC 720

Qy   2708 atatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggctgc 2767
          |||||||||| |||||||||||| |||     ||   ||| ||||||||||||||| |
Db    721 ATATGAGTGACGAAGAGGCGTTAAAATGCTCCA--GTTCCGATGTTTGACATGGGCCGCG 778

Qy   2768 ggaaacagtatcggcca 2784
          |||| ||||   ||| |
Db    779 GGAACCAGTTCGGGCAA 795


RESULT 12
BM014940
LOCUS       BM014940                832 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603640877F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5416818 5',
            mRNA sequence.
ACCESSION   BM014940
VERSION     BM014940.1  GI:16529294
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 832)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12063  row: j  column: 19
            High quality sequence stop: 811.
FEATURES             Location/Qualifiers
     source          1. .832
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5416818"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      231 a    205 c    186 g    208 t      2 others
ORIGIN


  Query Match             19.7%;  Score 701;  DB 10;  Length 832;
  Best Local Similarity   97.5%;  Pred. No. 3.1e-164;
  Matches  785;  Conservative    0;  Mismatches   12;  Indels    8;  Gaps    7;

Qy   2405 cacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacacc 2464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     29 CACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACC 88

Qy   2465 agtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctga 2524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     89 AGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGA 148

Qy   2525 cttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatact 2584
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    149 CTTCCCAGCAGCATGCGATTCTTA-TGACCTTGGGCGAACCTTTCCTACACACCCATACT 207

Qy   2585 tctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcac 2644
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 TCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCAC 267

Qy   2645 ttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttc 2704
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 TTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTNTGCTTC 327

Qy   2705 ttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggc 2764
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 TTCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGC 387

Qy   2765 tgcggaaacagtatcggccagacatgattatttttacagatccagatgtaccagctctcga 2824
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 TGCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGA 447

Qy   2825 ggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggcc 2884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 GGTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCC 507

Qy   2885 ccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctgggat 2944
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 CCAGCCTCTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGAT 567

Qy   2945 tcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaa-agtg 3003
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||| |
Db    568 TCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAACAGTG 627

Qy   3004 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 3063
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    628 GCTTTAAGTCTGNTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACC 687

Qy   3064 atagttgactttat-aaaaagcacgctacccaaccttggc-ttggtacagat-ggaaaag 3120
           |||||||||||||| |||||||||||| |||| |||||| |||||||||| ||||||| 
Db    688 ATAGTTGACTTTATAAAAAAGCACGCTAACCAAGCTTGGCTTTGGTACAGATGGGAAAAG 747

Qy   3121 accatcaatcaggtatttgaaatggacatcgctaaa--cagttacaagcttatgaagttg 3178
           |||||||||||||||||| |||||||||||| ||||     |   ||||||||||||||
Db    748 ACCATCAATCAGGTATTGGAAATGGACATCGGTAAAACAGTTACAAAGCTTATGAAGTTG 807

Qy   3179 agtacc-acgtccttcaagaagaac 3202
           |||||| ||||||||||||||||||
Db    808 AGTACCAACGTCCTTCAAGAAGAAC 832


RESULT  13
AL570425
LOCUS       AL570425                851 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL570425 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI020YF02 5
            prime, mRNA sequence.
ACCESSION   AL570425
VERSION     AL570425.1  GI:12926720
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 851)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .851
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI020YF02"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      240 a    214 c    220 g    174 t      3 others
ORIGIN


  Query Match             19.4%;  Score 689.8;  DB 9;  Length 851;
  Best Local Similarity   98.9%;  Pred. No. 2e-161;
  Matches  714;  Conservative    1;  Mismatches    4;  Indels    3;  Gaps    2;

Qy   2836 gattaccacagagacctctacaat--cacctggaggagcacgagatcggccccagcctct 2893
           |||||||||||||||||||||||     | ||||||||||||||||||||||||||||||
Db      1 GATTACCACAGAGACCTCTACAANTCACCTTGGAGGAGCACGAGATCGGCCCCAGCCTCT 60

Qy   2894 acgctgccccctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagcca 2953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCA 120

Qy   2954 gagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtc 3013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTC 180
```

```
Qy   3014 tgttgggaagccataagcccttgattctgcagcatgaaaacctagaaaccatagttgact 3073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACT 240

Qy   3074 ttataaaaagcacgctacccaaccttggcttggtacagatggaaaagaccatcaatcagg 3133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGG 300

Qy   3134 tatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttc 3193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TATTTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTC 360

Qy   3194 aagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagaga 3253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGA 420

Qy   3254 aaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaa 3313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAACCAACAGCAGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAA 480

Qy   3314 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 3373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGGTAGGATCCAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGC 540

Qy   3374 tgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggagg 3433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TGAAGCAGGCCATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGG 600

Qy   3434 agctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagc 3493
          |||||||||||||||  :|||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGCTGCGGCGGCGGA-SGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGC 659

Qy   3494 ccacgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcc 3553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 CCACGGGCGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCC 719

Qy   3554 tt 3555
          ||
Db    720 TT 721
```

```
RESULT  14
BE300005
LOCUS       BE300005                754 bp    mRNA    linear   EST 20-JUL-2000
DEFINITION  600944594F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960783 5',
            mRNA sequence.
ACCESSION   BE300005
VERSION     BE300005.1  GI:9183753
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM54  row: k  column: 24
            High quality sequence start: 2
            High quality sequence stop: 753.
FEATURES             Location/Qualifiers
     source          1. .754
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2960783"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      227 a    186 c    177 g    164 t
ORIGIN
```

```
  Query Match             19.2%;  Score 683.4;  DB 10;  Length 754;
  Best Local Similarity   99.1%;  Pred. No. 7.5e-160;
  Matches  750;  Conservative    0;  Mismatches    1;  Indels    6;  Gaps    6;

Qy   1878 ctcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatct 1937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATC- 59

Qy   1938 tggtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatatt 1997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 TGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATT 119

Qy   1998 cctccgagtagccacccccgcagaaggcgtgcgattcttccagcagatatgaagattattc 2057
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTC 179

Qy   2058 agagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgggccctt 2117
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||| |
Db    180 AGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAG-TTGTGAAGATGGGCCC-T 237

Qy   2118 tggccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaa 2177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    238 TGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAA 297

Qy   2178 ggctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaagc 2237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    298 GGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGC 357

Qy   2238 ctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccctg 2297
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    358 CTCTGAAAATGA-TTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCCTG 416

Qy   2298 tcttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaaat 2357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 TCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAAAT 476

Qy   2358 taagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcaccg 2417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 TAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCATCACCG 536

Qy   2418 aggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcccagcaa 2477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537 AGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGCAA 596

Qy   2478 acagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcagca 2537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    597 ACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAGCA 656

Qy   2538 tgcgattcttattgaccttgggcgaacctttcctacacacccatacttctctgcccagct 2597
          |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    657 TGCGATTCTTA-TGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAGCT 715

Qy   2598 tggagcaggacagctatcgctttacaacattttgaag 2634
          |||||||||||||||||  |||||||||||||| |
Db    716 TGGAGCAGGACAGCTATCGC-TTACAACATTTTGAGG 751
```

RESULT 15

```
BE561889
LOCUS       BE561889               892 bp    mRNA    linear   EST 15-AUG-2000
DEFINITION  601346094F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3678887 5',
            mRNA sequence.
ACCESSION   BE561889
VERSION     BE561889.1  GI:9805609
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM356  row: 1  column: 24
            High quality sequence start: 50
            High quality sequence stop: 758.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3678887"
                     /clone_lib="NIH_MGC_8"
                     /tissue_type="Burkitt lymphoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      224 a    240 c    242 g    186 t
ORIGIN
```

Query Match 19.2%; Score 682.6; DB 10; Length 892;
Best Local Similarity 95.3%; Pred. No. 1.3e-159;
Matches 790; Conservative 0; Mismatches 29; Indels 10; Gaps 8;

```
Qy   1628 aagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaagctcc 1687
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     63 AAGAGCCATCTGTGTGTGAAAAGGAGG-CTTGCCCATCTCTGAGAGCTCCTTTAAGCTCC 121

Qy   1688 tcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagccagctc 1747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 TCGGCTCCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTC 181

Qy   1748 cgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcactt-ccccatc 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    182 CGCTGTCGCCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCCATC 241

Qy   1807 gaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaactt 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 GAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTT 301

Qy   1867 atgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaa 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 ATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAA 361

Qy   1927 gcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctggagg 1986
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 GCAAATTATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGG 421

Qy   1987 cagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatat 2046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 CAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATAT 481

Qy   2047 gaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgtgaa 2106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 GAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAA 541

Qy   2107 gatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctccgagag 2166
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    542 GATGGGCCCTTTGGCCCCCCACCAGAGGAAACGAAAAGGACATCTCGTGAGCTCCGAGAG 601

Qy   2167 ctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaatcag 2226
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    602 CTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGCGAAGGAAAATCAG 661

Qy   2227 aagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaa 2286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 AAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAA 721

Qy   2287 attactccctg-tcttaaagaagtaactacagtgtggga--aaagatgcttagcactcca 2343
          |||||||||||| |||||||||||||||||||||||||  ||    ||||||||||||||
Db    722 ATTACTCCCTGTTCTTAAAGAAGTAACTACAGTGTGGGACCAAGATGGCTTAGCACTCCA 781

Qy   2344 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 2403
          ||||||||    ||| |||||| | |||  | ||||| ||||| |||||||| ||| | |
Db    782 GGAAGATC-CCCATTTAGTTTGGCCTGGTGACAATGC-CTCGGGTGTTGGGCCAGGGGGG 839

Qy   2404 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattcc 2452
          || |||||||||| | |||||   ||  | || |||||| ||||||||
Db    840 CC-CGTCATCACCGCGCTGAAACTGGG--AGTTTTAGCTGGGCAATTCC 885
```

Search completed: September 9, 2002, 16:41:28
Job time: 9256 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 19:18:41 ; Search time 706.85 Seconds
(without alignments)
8634.973 Million cell updates/sec

Title: US-09-762-311-4_COPY_176_3730
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3555	100.0	3983	21	AAZ50906	Human TBC-1 cDNA f
2	3555	100.0	3988	21	AAZ50907	Human TBC-1 cDNA f
3	3026.4	85.1	3326	22	AAI57854	Human polynucleoti
4	2940.2	82.7	3727	21	AAC77404	Human ORFX ORF2959
5	2833.8	79.7	3281	22	AAI59640	Human polynucleoti
6	2510.4	70.6	4039	19	AAV05886	Human Tbc-1 gene.
7	2112	59.4	2362	22	AAH18401	Human cDNA sequenc
8	1000.8	28.2	1422	22	AAS34787	cDNA encoding nove
9	999.4	28.1	2989	22	AAK94796	Human full-length
10	943.2	26.5	1269	22	AAS34785	cDNA encoding nove
11	785	22.1	883	22	AAH04450	Human cDNA clone (
12	574.2	16.2	849	22	AAK92006	Human cDNA 5'-end
13	574.2	16.2	849	22	AAK93666	Human cDNA clone r
14	422.6	11.9	17590	21	AAZ50904	Human TBC-1 partia
15	362	10.2	696	20	AAZ15229	Human gene express
16	306.2	8.6	5472	23	ABL02013	Drosophila melanog
17	284.4	8.0	99960	21	AAZ50905	Human TBC-1 partia
18	280	7.9	458	20	AAX40323	Human secreted pro
c 19	245.6	6.9	8105	23	ABL02012	Drosophila melanog
c 20	228.4	6.4	555	22	AAH09958	Human cDNA clone (
21	132.8	3.7	3691	22	AAF58616	Human RECAP polynu
22	132.8	3.7	4827	22	AAI59592	Human polynucleoti
23	132.8	3.7	5294	22	AAI57806	Human polynucleoti
24	128.8	3.6	3585	23	ABL12337	Drosophila melanog
25	123.8	3.5	154	23	AAS58636	cDNA #1312 encodin
26	119.4	3.4	2626	23	ABL18133	Drosophila melanog
27	119.4	3.4	7829	23	ABL18132	Drosophila melanog
28	108	3.0	2424	23	AAS71793	DNA encoding novel
29	83.8	2.4	1601	21	AAC33401	Arabidopsis thalia
c 30	79.6	2.2	6765	23	ABL12336	Drosophila melanog
31	78.4	2.2	300	20	AAZ13067	Human gene express
32	75.2	2.1	1492	21	AAC38750	Arabidopsis thalia
c 33	73.8	2.1	1140	21	AAZ97020	Human secreted pro
34	68.4	1.9	3342	22	AAF32789	Human secreted pro
35	68	1.9	2492	21	AAC78199	Human cancer assoc
36	65.2	1.8	1831	22	AAH64964	Human secreted pro
37	65	1.8	1117	22	AAI61302	Human polynucleoti
38	65	1.8	2035	22	AAF59600	Human cell cycle a
39	65	1.8	2482	22	AAI59516	Human polynucleoti
40	65	1.8	2666	22	AAH17866	Human cDNA sequenc
41	62.6	1.8	2531	23	ABL07359	Drosophila melanog
42	62.6	1.8	4661	23	ABL07358	Drosophila melanog
43	61.6	1.7	2801	22	AAF32643	Human cDNA encodin
c 44	61.6	1.7	2803	22	AAS27080	cDNA encoding nove
45	61.6	1.7	2903	22	AAK51641	Human polynucleoti

ALIGNMENTS

RESULT 1
AAZ50906
ID AAZ50906 standard; cDNA; 3983 BP.
XX
AC AAZ50906;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human TBC-1 cDNA from first transcript.
XX
KW TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;
KW Single nucleotide polymorphism; tissue differentiation; prostate cancer;
KW linkage analysis; genetic map; detection; diagnosis; genotyping;
KW transgenic animal; screening; alternative splicing; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..170
FT /*tag= a
FT /note= "Spans through exon 1 and part of exon 2"
FT CDS 171..3677
FT /*tag= b
FT /product= "Human TBC-1 protein"
FT /note= "First transcription product"
FT 3'UTR 3726..3983
FT /*tag= c
FT polyA_signal 3942..3947
FT /*tag= d
XX
PN WO200008209-A2.
XX

PD 17-FEB-2000.
XX
PF 06-AUG-1999; 99WO-IB01444.
XX
PR 07-AUG-1998; 98US-0095653.
XX
PA (GEST) GENSET.
XX
PI Blumenfeld M, Bougueleret L, Chumakov I;
XX
DR WPI; 2000-205736/18.
DR P-PSDB; AAY45096.
XX
PT New isolated human TBC-1 nucleic acids, useful for developing products
PT for the diagnosis and treatment of disorders involving cell
PT proliferation, particularly prostate cancer -
XX
PS Claim 3; Page 150-155; 166pp; English.
XX
CC The present sequence is the cDNA encoding the first transcription
CC product of human TBC-1 protein, comprising exons 1-2 and A-L.
CC Alternative splicing events result in two mRNA molecules, due to
CC splicing at two distinct first exons, exon1 and exon 1bis. TBC-1 gene
CC is mapped to a candidate region of prostate cancer on chromosome 4.
CC TBC-1 gene is involved in the regulation of cell cycle and tissue
CC differentiation in mammals. An alteration of TBC-1 sequence may be
CC associated with a pathological condition, resulting in abnormal cell
CC proliferation leading to cancer, e.g. prostate cancer. The biallelic
CC markers can be used for generation of genetic maps, linkage analysis and
CC association studies. TBC-1 sequence can be used for detection,
CC diagnosis, genotyping, production of transgenic animals and screening
CC of compounds for use in therapy.
XX
SQ Sequence 3983 BP; 1090 A; 1007 C; 1007 G; 878 T; 1 other;

Query Match 100.0%; Score 3555; DB 21; Length 3983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 230

Qy     61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    231 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 290

Qy    121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    291 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 350

Qy    181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    351 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 410

Qy    241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    411 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 470

Qy    301 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    471 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 530

Qy    361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    531 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 590

Qy    421 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    591 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 650

Qy    481 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    651 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 710

Qy    541 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    711 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 770

Qy    601 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    771 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 830

Qy    661 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    831 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 890

Qy    721 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    891 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 950

Qy    781 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    951 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 1010

Qy    841 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 1070

Qy    901 tacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttttaaggagatatcc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 tacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttttaaggagatatcc 1130

Qy    961 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1190

Qy   1021 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1250

Qy   1081 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1310

Qy   1141 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1370

Qy   1201 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1430

Qy   1261 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1490

Qy   1321 gagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaacag 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 gagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaacag 1550

Qy   1381 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1610

Qy   1441 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1670

Qy   1501 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1730

Qy   1561 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1731 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1790

Qy   1621 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1850

Qy   1681 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1910

Qy   1741 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1970

Qy   1801 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 2030

Qy   1861 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 2090

Qy   1921 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2091 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 2150

Qy   1981 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2151 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2210

Qy   2041 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2211 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2270

Qy   2101 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2271 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2330

Qy   2161 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2331 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2390

Qy   2221 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2391 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2450

Qy   2281 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2451 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2510

Qy   2341 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2511 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2570

Qy   2401 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2571 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2630

Qy   2461 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2631 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2690

Qy   2521 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2691 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2750

Qy   2581 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2751 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2810

Qy   2641 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2811 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2870

Qy   2701 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2871 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2930

Qy   2761 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2990

Qy   2821 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 3050

Qy   2881 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 3110

Qy   2941 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3170

Qy   3001 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3171 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3230

Qy   3061 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3290

Qy   3121 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3350

Qy   3181 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3410

Qy   3241 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3411 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3470

Qy   3301 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3471 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3530

Qy   3361 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3531 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3590

Qy   3421 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3591 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3650

Qy   3481 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3651 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3710

Qy   3541 cactgtccaggcctt 3555
          |||||||||||||||
Db   3711 cactgtccaggcctt 3725
```

RESULT 2
AAZ50907
ID AAZ50907 standard; cDNA; 3988 BP.
XX
AC AAZ50907;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human TBC-1 cDNA from second transcript.
XX
KW TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;

```
KW   Single nucleotide polymorphism; tissue differentiation; prostate cancer;
KW   linkage analysis; genetic map; detection; diagnosis; genotyping;
KW   transgenic animal; screening; alternative splicing; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..175
FT                   /*tag= a
FT   misc_feature    1..458
FT                   /*tag= b
FT                   /note= "Homologous to 5' EST obtained from human
FT                   pancreas cDNA library"
FT   CDS             176..3682
FT                   /*tag= c
FT                   /product= "Human TBC-1 protein"
FT                   /note= "Second transcription product"
FT   3'UTR           3731..3988
FT                   /*tag= d
FT   polyA_signal    3947..3952
FT                   /*tag= e
XX
PN   WO200008209-A2.
XX
PD   17-FEB-2000.
XX
PF   06-AUG-1999;   99WO-IB01444.
XX
PR   07-AUG-1998;   98US-0095653.
XX
PA   (GEST ) GENSET.
XX
PI   Blumenfeld M,  Bougueleret L,  Chumakov I;
XX
DR   WPI; 2000-205736/18.
DR   P-PSDB; AAY45096.
XX
PT   New isolated human TBC-1 nucleic acids, useful for developing products
PT   for the diagnosis and treatment of disorders involving cell
PT   proliferation, particularly prostate cancer  -
XX
PS   Claim 4; Page 156-161; 166pp; English.
XX
CC   The present sequence is the cDNA encoding the second transcription
CC   product of human TBC-1 protein, comprising exons 1bis to 2 and A-L.
CC   Alternative splicing events result in two mRNA molecules, due to
CC   splicing at two distinct first exons, exon1 and exon 1bis. TBC-1 gene
CC   is mapped to a candidate region of prostate cancer on chromosome 4.
CC   TBC-1 gene is involved in the regulation of cell cycle and tissue
CC   differentiation in mammals. An alteration of TBC-1 sequence may be
CC   associated with a pathological condition, resulting in abnormal cell
CC   proliferation leading to cancer, e.g. prostate cancer. The biallelic
CC   markers can be used for generation of genetic maps, linkage analysis and
CC   association studies. TBC-1 sequence can be used for detection,
CC   diagnosis, genotyping, production of transgenic animals and screening
CC   of compounds for use in therapy.
XX
SQ   Sequence 3988 BP; 1107 A; 983 C; 1002 G; 896 T; 0 other;
```

```
  Query Match             100.0%;  Score 3555;  DB 21;  Length 3988;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3555;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 235

Qy     61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    236 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 295

Qy    121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    296 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 355

Qy    181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    356 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 415

Qy    241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    416 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 475

Qy    301 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    476 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 535

Qy    361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    536 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 595

Qy    421 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    596 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 655

Qy    481 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    656 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 715

Qy    541 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    716 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 775

Qy    601 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    776 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 835

Qy    661 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    836 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 895

Qy    721 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    896 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 955

Qy    781 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    956 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 1015

Qy    841 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1016 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 1075

Qy    901 tacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttttaaggagatatcc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1076 tacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttttaaggagatatcc 1135

Qy    961 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1136 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1195

Qy   1021 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1196 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1255

Qy   1081 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1256 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1315

Qy   1141 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1316 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1375

Qy   1201 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1376 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1435
Qy  1261 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1436 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1495
Qy  1321 gagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaacag 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1496 gagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaacag 1555
Qy  1381 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1556 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1615
Qy  1441 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1616 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1675
Qy  1501 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1676 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1735
Qy  1561 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1736 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1795
Qy  1621 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1796 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1855
Qy  1681 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1856 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1915
Qy  1741 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1916 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1975
Qy  1801 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1976 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 2035
Qy  1861 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2036 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 2095
Qy  1921 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2096 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 2155
Qy  1981 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2156 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2215
Qy  2041 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2216 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2275
Qy  2101 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2276 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2335
Qy  2161 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2336 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2395
Qy  2221 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2396 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2455
Qy  2281 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2456 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2515
```

```
Qy  2341 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2516 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2575
Qy  2401 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2576 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2635
Qy  2461 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2636 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2695
Qy  2521 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2696 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2755
Qy  2581 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2756 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2815
Qy  2641 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2816 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2875
Qy  2701 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2876 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2935
Qy  2761 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2936 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2995
Qy  2821 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2996 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 3055
Qy  2881 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3056 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 3115
Qy  2941 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3116 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3175
Qy  3001 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3176 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3235
Qy  3061 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3236 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3295
Qy  3121 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3296 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3355
Qy  3181 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3356 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3415
Qy  3241 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3416 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3475
Qy  3301 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3476 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3535
Qy  3361 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3536 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3595
```

```
Qy   3421 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3596 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3655

Qy   3481 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3656 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3715

Qy   3541 cactgtccaggcctt 3555
          |||||||||||||||
Db   3716 cactgtccaggcctt 3730
```

RESULT 3
AAI57854
ID AAI57854 standard; cDNA; 3326 BP.
XX
AC AAI57854;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 57.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0633450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM38698.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 57; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 3326 BP; 924 A; 824 C; 843 G; 735 T; 0 other;

Query Match 85.1%; Score 3026.4; DB 22; Length 3326;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3030; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy    520 gaggtgctcttctgcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatc 579
          | | | ||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db     43 gggatcctctagagcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatc 102

Qy    580 gacgagtgcatcgagaagttcaatcacgtcagcggcagccgggggtccgagagccccccgc 639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103 gacgagtgcatcgagaagttcaatcacgtcagcggcagccgggggtccgagagccccccgc 162

Qy    640 cccaacccgccccatgccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgc 699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    163 cccaacccgccccatgccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgc 222

Qy    700 aagtccttctcccagcccggcctgcgctcgctggcctttaggaaggagctgcaggatggg 759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    223 aagtccttctcccagcccggcctgcgctcgctggcctttaggaaggagctgcaggatggg 282

Qy    760 ggcctccgaagcagcggcttcttcagctccttcgaggagagcgacattgagaaccacctc 819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    283 ggcctccgaagcagcggcttcttcagctccttcgaggagagcgacattgagaaccacctc 342

Qy    820 attagcggacacaatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttc 879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    343 attagcggacacaatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttc 402

Qy    880 acgattggccagtctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggag 939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    403 acgattggccagtctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggag 462

Qy    940 aaaaattttaaggagatatccttttgctctcagggcatcagacacgtggaccactttggg 999
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    463 aaaaattttaaggagatatccttttgctctcagggcatcagacacgtggaccactttggg 522

Qy   1000 tttatctgtcgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcag 1059
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    523 tttatctgtcgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcag 582

Qy   1060 tgcacaaatgaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtg 1119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    583 tgcacaaatgaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtg 642

Qy   1120 gccgcagtgcagcagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaa 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    643 gccgcagtgcagcagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaa 702

Qy   1180 agcctgcacaagctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaa 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    703 agcctgcacaagctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaa 762

Qy   1240 ctgcaaaagcacctgacgacattaaccaatcaggagcaggcgactatttttgaagaggtt 1299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    763 ctgcaaaagcacctgacgacattaaccaatcaggagcaggcgactatttttgaagaggtt 822

Qy   1300 cagaaattgagaccgagaaatgagcagcgagagaatgaattgattatttcttttctgaga 1359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    823 cagaaattgagaccgagaaatgagcagcgagagaatgaattgattatttcttttctgaga 882

Qy   1360 tgtttatatgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcg 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    883 tgtttatatgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcg 942

Qy   1420 cagatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggcta 1479
```

```
Db    943 cagatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggcta 1002
Qy   1480 gatatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcc 1539
Db   1003 gatatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcc 1062
Qy   1540 cggggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctcc 1599
Db   1063 cggggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctcc 1122
Qy   1600 ctgtctagtacattaagtaaccaccagcaaagagccatctgtgtgtgaaaaggaggccttg 1659
Db   1123 ctgtctagtacattaagtaaccaccagcaaagagccatctgtgtgtgaaaaggaggccttg 1182
Qy   1660 cccatctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcg 1719
Db   1183 cccatctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcg 1242
Qy   1720 gagagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcga 1779
Db   1243 gagagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcga 1302
Qy   1780 gcaaacaccctgagtcacttccccatcgaatgccaggaacctccacaacctgcccggggg 1839
Db   1303 gcaaacaccctgagtcacttccccatcgaatgccaggaacctccacaacctgcccggggg 1362
Qy   1840 tccccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcct 1899
Db   1363 tccccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcct 1422
Qy   1900 catgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcct 1959
Db   1423 catgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcct 1482
Qy   1960 gtgaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcag 2019
Db   1483 gtgaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcag 1542
Qy   2020 aaggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttccccca 2079
Db   1543 aaggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttccccca 1602
Qy   2080 cgatctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaaag 2139
Db   1603 cgatctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaaag 1662
Qy   2140 aaaaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg 2199
Db   1663 aaaaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg 1722
Qy   2200 ctgcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaac 2259
Db   1723 ctgcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaac 1782
Qy   2260 aagcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtg 2319
Db   1783 aagcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtg 1842
Qy   2320 tgggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatg 2379
Db   1843 tgggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatg 1902
Qy   2380 cactcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttcta 2439
Db   1903 cactcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttcta 1962
Qy   2440 gctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgcca 2499
Db   1963 gctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgcca 2022
Qy   2500 tacaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttggg 2559
```

```
Db   2023 tacaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttggg 2082
Qy   2560 cgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctt 2619
Db   2083 cgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctt 2142
Qy   2620 tacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctc 2679
Db   2143 tacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctc 2202
Qy   2680 agctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctc 2739
Db   2203 agctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctc 2262
Qy   2740 aagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattatttta 2799
Db   2263 aagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattatttta 2322
Qy   2800 cagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaat 2859
Db   2323 cagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaat 2382
Qy   2860 cacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatg 2919
Db   2383 cacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatg 2442
Qy   2920 tttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcag 2979
Db   2443 tttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcag 2502
Qy   2980 ggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgatt 3039
Db   2503 ggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgatt 2562
Qy   3040 ctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccctt 3099
Db   2563 ctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccctt 2622
Qy   3100 ggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacag 3159
Db   2623 ggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacag 2682
Qy   3160 ttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcct 3219
Db   2683 ttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcct 2742
Qy   3220 ctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacag 3279
Db   2743 ctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacag 2802
Qy   3280 aaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcc 3339
Db   2803 aaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcc 2862
Qy   3340 accattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaa 3399
Db   2863 accattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaa 2922
Qy   3400 ctggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagccc 3459
Db   2923 ctggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagccc 2982
Qy   3460 agcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagg 3519
Db   2983 agcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagg 3042
Qy   3520 agagattgcaacaccatcccacactgtccaggcctt 3555
Db   3043 agagattgcaacaccatcccacactgtccaggcctt 3078
```

RESULT 4
AAC77404

ID AAC77404 standard; cDNA; 3727 BP.
XX
AC AAC77404;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2959 polynucleotide sequence SEQ ID NO:5917.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB43195.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 5; Page 5094-5096; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 3727 BP; 1031 A; 889 C; 931 G; 875 T; 1 other;

Query Match 82.7%; Score 2940.2; DB 21; Length 3727;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 3005; Conservative 0; Mismatches 13; Indels 7; Gaps 5;

```
Qy   532 tgcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatc 591
Db     4 tgcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatc 63

Qy   592 gagaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccc 651
Db    64 gagaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccc 123

Qy   652 catgccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcc 711
Db   124 catgccgcgcgccacagggagccaggacctgtgcgcaggcccatgcgcaagtccttctcc 183

Qy   712 cagcccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagc 771
Db   184 cagcccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagc 243

Qy   772 agcggcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacac 831
Db   244 agcggcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacac 303

Qy   832 aatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccag 891
Db   304 aatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccag 363

Qy   892 tctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttaag 951
Db   364 tctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttaag 423

Qy   952 gagatatccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgg 1011
Db   424 gagatatccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgg 483

Qy  1012 gagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
Db   484 gagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 543

Qy  1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
Db   544 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 603

Qy  1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
Db   604 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 663

Qy  1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
Db   664 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 723

Qy  1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
Db   724 ctgacgacattaaccaatcaggagcaggcgactattttgaagaggttcagaaattgaga 783

Qy  1312 ccgagaaatgagcagcgagagaatgaattgattatttctttttctgagatgtttatatgaa 1371
Db   784 ccgagaaatgagcagcgagagaatgaattgattatttctttttctgagatgtttatatgaa 843

Qy  1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
Db   844 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 903

Qy  1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
Db   904 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 963

Qy  1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1551
Db   964 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1023
```

```
Qy   1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1083

Qy   1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1084 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1143

Qy   1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1203

Qy   1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1204 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1263

Qy   1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1323

Qy   1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1383

Qy   1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1443

Qy   1972 aggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgat 2031
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 aggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgat 1503

Qy   2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 1563

Qy   2092 gaaccagtttgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatct 2151
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   1564 gaaccagtttgtgaagatgggccctttgg-cccccaccagaggaaaagaaaaggacatct 1622

Qy   2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgc-tgcttagaat 2210
          |||||||||||||||||||||||| |||||||||||||||||||||| |||||||||||
Db   1623 cgtgagctccgagagctgtggcaaagggctattcttcaacagatactgcntgcttagaat 1682

Qy   2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1683 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 1742

Qy   2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1743 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 1802

Qy   2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1803 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 1862

Qy   2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1863 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 1922

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1923 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 1982

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1983 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2042

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2043 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2102

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2103 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2162

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2163 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2222

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2223 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2282

Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2283 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2342

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2343 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2402

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2403 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2462

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2463 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 2522

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2523 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 2582

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2583 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 2642

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2643 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 2702

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2703 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 2762

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2763 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 2822

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2823 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 2882

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
          |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2883 g--cctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 2940

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaa 3530
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2941 gcctgagtgcacgcagcccgagcccacgggcgactgacagct-tgcaggagagattgcaa 2999

Qy   3531 caccatcccacactgtccaggcctt 3555
          |||||| ||||||||||||||||
Db   3000 caccat--cacactgtccaggcctt 3022
```

```
RESULT   5
AAI59640
ID   AAI59640 standard; cDNA; 3281 BP.
XX
AC   AAI59640;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 3629.
```

```
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM40484.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 3629; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3281 BP; 912 A; 812 C; 824 G; 733 T; 0 other;


  Query Match             79.7%;  Score 2833.8;  DB 22;  Length 3281;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2983;  Conservative    0;  Mismatches   32;  Indels   20;  Gaps   11;

Qy    541 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 60

Qy    601 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 120

Qy    661 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    121 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 180

Qy    721 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 240

Qy    781 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 300

Qy    841 cagcccacagatat--cgaggaaaatcgaactat---gctcttcacgattggccagtctg 895
          ||||||||||||||  | ||| |||||||||  || |·  | ||||||||||
Db    301 cagcccacagatattcgaggaaaattcgaactattgctctttaacgattgggccagtctg 360

Qy    896 aagtttacctcatcagtcctgacaccaaaaaaat-----agcattggagaaaaattttaa 950
          |||||||||||||||||||||||||||||||||     |      || |||  ||| |
Db    361 aagtttacctcatcagtcctgacaccaaaaaaattaggcattgggagaaaaattttaag 420

Qy    951 ggagatatccttttgctctcagggcatcagaca-cgtggaccactttggg-tttatctgt 1008
          |||||||||||||||||||||||||||||||| |||||||||||| || |||||||||
Db    421 ggagatatccttttgctctcagggcatcagacaccgtggaccactttggttttatctgt 480

Qy   1009 cgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaat 1068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 cgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaat 540

Qy   1069 gaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtg 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtg 600

Qy   1129 cagcagacagctaaggcgccagcccagctgtgtgagggctg--ccccctgcaaagcc--t 1184
          ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db    601 cagcagacagctaaggcgccagcccagctgtgtgagggctgccccccctgcaaagccctg 660

Qy   1185 gcacaagctctgtgagaggatagagggaatgaatt-cttccaaaacaaaactagaactgc 1243
          |||||||||||||||||||||||||||||||||| || ||||||||||||||||||||||
Db    661 ccacaagctctgtgagaggatagagggaatgaattcctcccaaaacaaaactagaactgc 720

Qy   1244 aaaagca-cctgacgacattaa-ccaatcaggagcaggcgactatttttgaagaggttca 1301
          ||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    721 aaaagcaccctgacgacattaacccaatcaggagcaggcgactatttttgaagaggttca 780

Qy   1302 gaaattgagaccgagaaatgagcagcgagagaatgaattgattatttctttttctgagatg 1361
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    781 gaaattgagaccgagaaatgagcagcgagagaatgaattgattatttctttttttgagatg 840

Qy   1362 tttatatgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcgca 1421
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 tttatttgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcgca 900

Qy   1422 gatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggctaga 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 gatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggctaga 960

Qy   1482 tatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccg 1541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 tatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccg 1020

Qy   1542 gggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctccct 1601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctccct 1080

Qy   1602 gtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcc 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 gtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcc 1140

Qy   1662 catctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcgga 1721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 catctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcgga 1200

Qy   1722 gagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagc 1781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 gagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagc 1260
```

```
Qy   1782 aaacaccctgagtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtc 1841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 aaacaccctgagtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtc 1320

Qy   1842 cccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctca 1901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 cccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctca 1380

Qy   1902 tgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgt 1961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgt 1440

Qy   1962 gaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaa 2021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 gaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaa 1500

Qy   2022 ggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttcccccacg 2081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttcccccacg 1560

Qy   2082 atctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaa 2141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 atctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaa 1620

Qy   2142 aaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgct 2201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 aaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgct 1680

Qy   2202 gcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaa 2261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 gcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaa 1740

Qy   2262 gcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtg 2321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 gcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtg 1800

Qy   2322 ggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgca 2381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgca 1860

Qy   2382 ctcggctgttgggcaaggtgtg-ccacgtcatcaccgaggtgaaatctggaaatttctag 2440
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1861 ctcggctgttgggcaaggtgtgcccacgtcatcaccgaggtgaaatctggaaatttctag 1920

Qy   2441 ctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccat 2500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccat 1980

Qy   2501 acaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggc 2560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 acaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggc 2040

Qy   2561 gaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgcttt 2620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 gaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgcttt 2100

Qy   2621 acaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctca 2680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 acaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctca 2160

Qy   2681 gctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctca 2740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 gctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctca 2220

Qy   2741 agtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttac 2800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 agtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttac 2280

Qy   2801 agatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatc 2860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 agatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatc 2340

Qy   2861 acctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgt 2920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 acctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgt 2400

Qy   2921 ttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagg 2980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 ttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagg 2460

Qy   2981 gaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattc 3040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 gaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattc 2520

Qy   3041 tgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttg 3100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 tgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttg 2580

Qy   3101 gcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagt 3160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 gcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagt 2640

Qy   3161 tacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctc 3220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 tacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctc 2700

Qy   3221 tcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacaga 3280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 tcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacaga 2760

Qy   3281 accttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcca 3340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 accttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcca 2820

Qy   3341 ccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaac 3400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 ccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaac 2880

Qy   3401 tggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagccca 3460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2881 tggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcaaagccca 2940

Qy   3461 gcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagga 3520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 gcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagga 3000

Qy   3521 gagattgcaacaccatcccacactgtccaggcctt 3555
          |||||||||||||||||||||||||||||||||
Db   3001 gagattgcaacaccatcccacactgtccaggcctt 3035
```

RESULT 6
AAV05886
ID AAV05886 standard; DNA; 4039 BP.
XX
AC AAV05886;
XX
DT 01-JUN-1998 (first entry)
XX
DE Human Tbc-1 gene.
XX
KW tre-2; BUB2; cdc16; mast cell; probe; hybridisation; subtraction method;
KW mouse; transcription factor; differentiation; proliferation; human; ds;
KW acute myelogenous leukaemia.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 403..3828
FT /*tag= a
FT /product= Tbc1_protein
XX
PN US5700927-A.
XX
PD 23-DEC-1997.

```
XX
PF   23-DEC-1994;   94US-0363300.
XX
PR   23-DEC-1994;   94US-0363300.
XX
PA   (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI   Richardson P,  Zon L;
XX
DR   WPI; 1998-062437/06.
DR   P-PSDB; AAW44777.
XX
PT   DNA encoding Tbc1 polypeptide - useful for treating leukaemia
XX
PS   Claim 3; Fig 1A-B; 22pp; English.
XX
CC   This is the nucleotide sequence which encodes a Tbc1 (tre-2, BUB2 and
CC   cdc16) polypeptide, isolated from a bone marrow-derived mast cell DNA
CC   library.  The screen was carried out using a probe generated by a
CC   subtraction method which compared mRNA expression in an undifferentiated
CC   mast cell line P815 and cell line PGT6 (P815 cells transformed to express
CC   the murine GATA-1 transcription factor - a factor which controls the
CC   expression of genes involved in mast cell differentiation).  Tbc1 encodes
CC   a protein involved in the coupling of cell proliferation to cell
CC   differentiation, which can be used to treat leukaemia (especially acute
CC   myelogenous leukaemia) by causing leukaemic cells to differentiate.
XX
SQ   Sequence 4039 BP; 1042 A; 1096 C; 1060 G; 841 T; 0 other;


  Query Match             70.6%;  Score 2510.4;  DB 19;  Length 4039;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches 2949;  Conservative    0;  Mismatches  536;  Indels   37;  Gaps    8;

Qy      1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59
          |||||  |||| |||||||||| ||||| |||| | || |||||||| ||||| |||||
Db    292 atggaggcaatcacattcacagcgaggaagcatccgtttccctaacgaagtctccgtgga 351

Qy     60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
            || ||| ||||||||||  || ||||| | || |||||||| || ||||| |||||||||
Db    352 cttcggcttgcagctggttgggtctccttacccgtgcattctctcaccactatgcccatg 411

Qy    118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
          ||||| ||||| || || ||||| ||||||||||  | |||| | | | |||||| |||
Db    412 ctgccgtgggtggtagccgaggtacgaagactcagtggccagtgctctaaaaaggagcct 471

Qy    178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
             || |||||||||| ||||| ||||| |||||| | || || ||||| |
Db    472 aggacaaagcaagtccgcctttgggtttcgccctccggactgcggtgcgagcctgacctg 531

Qy    238 gggagaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcag 297
          | || ||||||||  ||||| || || |||| ||||||||||||||||||||||||||||
Db    532 gagaaaagtcaaccatgggacccgctcatctgttccagcatctttgagtgcaagcctcag 591

Qy    298 cgtgttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
          ||||| |||||||||||||||||||||| ||||| || |||||||||| || ||||||
Db    592 cgtgtgcacaaactgattcacaacagtcacgacccgagctactttgcttgcctcattaag 651

Qy    358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
          || || ||||  ||| ||||||| |||||||||||| ||||| || ||||||||||||
Db    652 gaggatgctgcgcacaggcagagcctctgctatgtgtttaaagcagacgatcaaacaaaa 711

Qy    418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
          |||||||||||||||||||||||||||| ||||| |||||||| ||||| ||||| ||||||
Db    712 gtgcctgagatcatcagctccatccggcaggccgggaagattgcccgccaggaagagctg 771

Qy    478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
          |  ||||| |||||||||||||| || ||  |||| ||||||||||||||||||||| |||
Db    772 cgttgcccctccgagttcgacgataccttcgccaaaaagttcgaggtgctcttctgtggc 831

Qy    538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
          || ||||| ||||| |||||||||| || || || || ||||| |||||||| |||||||||
Db    832 cgggtgactgtggctcacaagaaggccccacccgcactgattgacgagtgtatcgagaag 891

Qy    598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
          ||||| || || ||| |  | ||  |  | ||  |       |||| |  ||  |   |
Db    892 ttcaaccatgtgagctgtggtcgcagaacggactgggaagcgcccaccgggcagccatca 951

Qy    658 gcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagccc 717
          ||                  ||||| | ||||||||||||||| |||||||| |||||
Db    952 gc------------------gcctggccccaggcccatgcgcaaatccttctcacagcct 993

Qy    718 ggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggc 777
          || |||||||||||||||| ||||||||| | ||||| |   ||||||| || |||  |
Db    994 ggactgcgctcgctggccttcaggaaggagttccaggacgctagcctccgcagtagcacc 1053

Qy    778 ttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatatt 837
          ||         || |   || |  ||||| ||||||||||||  | || |||||| |
Db   1054 tt------tagctcctttgacaatgacatagagaaccacctcatcggtgggcacaatgtg 1107

Qy    838 gtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaa 897
          || |||||||||| || ||||| || |||||||||| |||||||||||||  ||||||
Db   1108 gttcagcccacagacatggaggagaaccgaactatgctgttcacgattggcccatctgaa 1167

Qy    898 gtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaattttaaggagata 957
          |||||||||||||||||||||||||||||| || ||| ||||||||||||||||||||||
Db   1168 gtttacctcatcagtcctgacaccaaaaagattgcactggagaaaaattttaaggagata 1227

Qy    958 tccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtct 1017
          |||||||||||||||||||||||||||||| |||||||||||||| || ||||| || ||||
Db   1228 tccttttgctctcagggcatcagacatgtggaccactttggattcatctgccgagagtgc 1287

Qy   1018 tc------cggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
          ||      |||  |||||||||||||||||||||||||||||||| ||||||||||||||
Db   1288 tcgggtggcggcagtggcggctttcattttgtctgttacgtgttccagtgcacaaatgaa 1347

Qy   1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
          |||||||||| || || |||||||| ||||| ||||| ||||||| || || ||||||
Db   1348 gctctggttgacgagatcatgatgactctgaagcaggctttcacggtagctgcggtgcag 1407

Qy   1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
          ||||| |||||||| |||||||||| |||||||||||||||| ||||| ||||||||||
Db   1408 cagacggctaaggcaccagcccagctctgtgagggctgcccctgcaaggcctgcacaag 1467

Qy   1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
          ||||| || |||||||||||||||||||| || ||||| ||| ||||||| || ||||||
Db   1468 ctctgcgaaaggatagagggaatgaattcatctaaaaccaaattagaactccagaagcac 1527

Qy   1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
            |||| ||| | |||||||||||||||||| || || || || ||||||||||||||||||
Db   1528 ttgaccacactgaccaatcaggagcaggccaccatattcgaggaggttcagaaattgaga 1587

Qy   1312 ccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaa 1371
          || ||||| |||||||||||||||||||||| ||||||||||||||| || |||||||||
Db   1588 ccaagaaacgagcagcgagagaatgaattaattatttcttttctgaggtgcttatatgaa 1647

Qy   1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
          ||||| || ||||| |||| ||| | ||||| |   |||||||||   ||| ||||||||
Db   1648 gagaagcaaaaagagcacagccacactggggcgccaaagcagacactacaggtggcagca 1707

Qy   1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
          |||||||||| |||||  | |||||||||| |  || || ||| |||||  ||||||
Db   1708 gagaatattgggagtgacctgccacccagtgctagccggttcaggttagattcgctgaag 1767

Qy   1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1551
          ||||  |||||||| || |||||||||  |||| || ||| ||||||||||||||||||
Db   1768 aacagagcaaagaggtccttaacagagtccctagagagcattctgtcccggggtaataaa 1827

Qy   1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
          |||||||||||||||| || |||  |||||||||||||| ||||||   ||||||||
Db   1828 gccagaggcctgcaggaccattccgccagtgtggatctggacagctccacttctagtact 1887

Qy   1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
            ||||||||||||||||||||  ||  || |||| |||||||||| ||| ||||||||
Db   1888 ctaagtaacaccagcaaagagctgtccatgggtgacaaggaggccttccccgtctctgag 1947
```

```
Qy   1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
Db   1948 acctccttcaagctccttggctcctcagatgacctgtccagtgactcagagggccacatt 2007

Qy   1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
Db   2008 gcagaagagtctgccctgttgtcaccccagcaggcgttcagaaggagagccaacaccctg 2067

Qy   1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
Db   2068 agtcatttcccagtagagtgccctgcgcctccagaacctgcccagagctctccaggggtc 2127

Qy   1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
Db   2128 tctcaaaggaaactcatgcggtaccactccgtgagcacagagacgcctcatgaacgcaag 2187

Qy   1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
Db   2188 gactttgaatccaaagcaaaccacctgggtgacacagatgggacccccgtgaagacccgg 2247

Qy   1972 aggcattcctggaggcagcagatattcctccgagtagccacccccgcagaaggcgtgcgat 2031
Db   2248 cggcactcgtggagacagcagatattccttcgagtggccactccacagaaggcttgtgac 2307

Qy   2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
Db   2308 tccccgagcagatatgaagattattccgagctgggagagctccctccacgctccccttta 2367

Qy   2092 gaaccagtttgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatct 2151
Db   2368 gaaccggtgtgtgaggacggcccatttgg-ccagtacaggaagaaaagaggaagacgtca 2426

Qy   2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg-ctgcttagaat 2210
Db   2427 cgcgagcttcgagagctgtggaaaaaggccatcttgcagcagatcctgcctcgtcaggat 2486

Qy   2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
Db   2487 ggagaaggagaatcagaagctacaagcctctgaaaacgatttgctgaacaaacgcctcaa 2546

Qy   2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
Db   2547 gcttgactatgaagaaatcactccgtgtcttaaagaagtcactacagtgtgggaaaagat 2606

Qy   2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
Db   2607 gcttagcactccaggaagatccaaaattaagtttgacatggaaaaagtgcactcagctgt 2666

Qy   2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
Db   2667 tgggcaaggtgtgccacgtcatcaccgaggtgagatctggaaatttctagctgagcagtt 2726

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
Db   2727 ccaccttaaacacccatttcctagtaaacagcagccaaaggacgtgccctacaaagagct 2786

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
Db   2787 cctgaagaagctgacctcgcagcagcacgccattctcatcgacctcgggcgaacctttcc 2846

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
Db   2847 aacacatccatacttctctgcccagcttggagcaggtcagctgtcactttacaacattct 2906

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
Db   2907 gaaggcctactcgcttctggaccaggaggttggatactgccaaggtctcagctttgtggc 2966

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
Db   2967 aggcattttgcttcttcacatgagtgaggaagaggcgttcaagatgctcaagttcctgat 3026

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
Db   3027 gtttgacatggggctgcggaaacagtatcggccagacatgattattttgcagatccagat 3086

Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
Db   3087 gtaccagctgtcacggctcctccacgattaccaccgagacctctacaaccacctggaaga 3146

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
Db   3147 gcacgagactggccccctacgtacgcggctccctggtttctcaccgtgttcgcctcaca 3206

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
Db   3207 gttcccactcggctttgtagccagagtctttgatatgatcttccttcagggatcagaggt 3266

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
Db   3267 catatttaaagtagctttaagtcttttggggagccataagcccttgattctacagcatga 3326

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
Db   3327 gaacctggaaaccatcgtggacttcataaagaacacactccccaacctgggcctggtgca 3386

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
Db   3387 gatggagaagaccatcagtcaggtgtttgagatggacatcgccaagcagctccaggccta 3446

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
Db   3447 tgaggtcgagtaccacgtcgtccaggaggagcttattgagtcctcgcctctcagtgacaa 3506

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
Db   3507 ccaaagaatggagaaattggagaaaaccaacagcacgttgcgcaaacagaaccttgacct 3566

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
Db   3567 cctggagcagttgcaggtggcaaatgctaggatccaaagccttgaagccacggtagagaa 3626

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
Db   3627 acttcttaccagcgagagtaagctgaagcagcgtgcgctgaccctggaggtggagcg--t 3684

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
Db   3685 cgccctgctgcagatggtggaggagctgcggaggcaaagcgcccggcccagcactccaga 3744

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
Db   3745 gccagactgcacccagctggagcccacaggcgattgaccgct 3786
```

```
RESULT   7
AAH18401
ID   AAH18401 standard; cDNA; 2362 BP.
XX
AC   AAH18401;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:18465.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
```

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18465; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2362 BP; 657 A; 581 C; 578 G; 546 T; 0 other;

Query Match 59.4%; Score 2112; DB 22; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1444 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 1503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 60

Qy   1504 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 1563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 120

Qy   1564 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 1623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 180

Qy   1624 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttttaag 1683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttttaag 240

Qy   1684 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 1743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 300

Qy   1744 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 1803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 360

Qy   1804 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 1863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 420

Qy   1864 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 1923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 480

Qy   1924 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 1983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 540

Qy   1984 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 2043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 600

Qy   2044 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 2103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 660

Qy   2104 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 2163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 720

Qy   2164 gagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaat 2223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaat 780

Qy   2224 cagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaa 2283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 cagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaa 840

Qy   2284 gaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactcca 2343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactcca 900

Qy   2344 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 2403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 960

Qy   2404 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacac 2463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacac 1020

Qy   2464 cagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctg 2523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 cagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctg 1080

Qy   2524 acttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatac 2583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 acttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatac 1140

Qy   2584 ttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactca 2643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactca 1200

Qy   2644 cttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgctt 2703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 cttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgctt 1260

Qy   2704 cttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatgggg 2763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 cttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatgggg 1320

Qy   2764 ctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 2823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 1380

Qy   2824 aggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggc 2883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 aggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggc 1440

Qy   2884 cccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctggga 2943
```

```
Db    1441 cccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctggga 1500

Qy    2944 ttcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtg 3003
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 ttcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtg 1560

Qy    3004 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 3063
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 1620

Qy    3064 atagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaagacc 3123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 atagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaagacc 1680

Qy    3124 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtac 3183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtac 1740

Qy    3184 cacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggat 3243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 cacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggat 1800

Qy    3244 aaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttg 3303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 aaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttg 1860

Qy    3304 caggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagcagt 3363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 caggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagcagt 1920

Qy    3364 gagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcag 3423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 gagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcag 1980

Qy    3424 acggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacg 3483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 acggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacg 2040

Qy    3484 cagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatcccacac 3543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 cagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatcccacac 2100

Qy    3544 tgtccaggcctt 3555
           ||||||||||||
Db    2101 tgtccaggcctt 2112


RESULT   8
AAS34787
ID   AAS34787 standard; cDNA; 1422 BP.
XX
AC   AAS34787;
XX
DT   04-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human neoplastic disease associated polypeptide #21.
XX
KW   Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW   hyperproliferative disorder; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155163-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01358.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
```

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465558/50.
DR P-PSDB; AAU21588.
XX
PT Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid
PT arthritis -
XX
PS Claim 4; SEQ ID No 31; 687pp; English.
XX
CC The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of disorders involving neoplastic disease such as
CC hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC also be useful for treating other disorders such as neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC and renal disorders. The polynucleotide sequences of the invention are
CC also useful in gene therapy. AAS34767-AAS35050 represent cDNA sequences
CC encoding for the novel human neoplastic disease associated polypeptides
CC of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1422 BP; 399 A; 322 C; 337 G; 364 T; 0 other;

Query Match 28.2%; Score 1000.8; DB 22; Length 1422;
Best Local Similarity 99.3%; Pred. No. 1.5e-262;
Matches 1005; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy   2544 tcttattgaccttgggcgaacctttcctacacacccatacttctctgcccagcttggagc 2603
          | || ||  |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 tattcttttcacagggcgaacctttcctacacacccatacttctctgcccagcttggagc 198

Qy   2604 aggacagctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtggg 2663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 aggacagctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtggg 258

Qy   2664 atattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaaga 2723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 atattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaaga 318

Qy   2724 ggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggcc 2783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 ggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggcc 378

Qy   2784 agacatgattattttacagatccagatgtaccagctctcgaggttgcttcatgattacca 2843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 agacatgattattttacagatccagatgtaccagctctcgaggttgcttcatgattacca 438

Qy   2844 cagagacctctacaatcacctggaggagcacgagatcggccccagcctctacgctgcccc 2903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    439 cagagacctctacaatcacctggaggagcacgagatcggccccagcctctacgctgcccc 498

Qy   2904 ctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccagagtctttga 2963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    499 ctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccagagtctttga 558

Qy   2964 tatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaag 3023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    559 tatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaag 618

Qy   3024 ccataagcccttgattctgcagcatgaaaacctagaaaccatagttgactttataaaaag 3083
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    619 ccataagcccttgattctgcagcatgaaaacctagaaaccatagttgactttataaaaag 678

Qy   3084 cacgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaat 3143
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    679 cacgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaat 738

Qy   3144 ggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaagaagaact 3203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    739 ggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaagaagaact 798
```

```
Qy   3204 tatcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacag 3263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 tatcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacag 858

Qy   3264 cagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggat 3323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 cagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggat 918

Qy   3324 ccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggc 3383
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 ccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggc 978

Qy   3384 catgcttaccttagaactggagcggtcggccctgctgcagacggtggaggagctgcggcg 3443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 catgcttaccttagaactggagcggtcggccctgctgcagacggtggaggagctgcggcg 1038

Qy   3444 gcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcga 3503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 gcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcga 1098

Qy   3504 ctgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 ctgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 1150


RESULT   9
AAK94796
ID   AAK94796 standard; cDNA; 2989 BP.
XX
AC   AAK94796;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human full-length cDNA, SEQ ID NO: 3914.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-0114089.
XX
PR   08-JUL-1999;   99JP-0194486.
PR   11-JAN-2000; 2000JP-0118774.
PR   02-MAY-2000; 2000JP-0183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
DR   WPI; 2001-524255/58.
DR   P-PSDB; AAM93840.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their
PT   use in genetic manipulation -
XX
PS   Claim 8; SEQ ID NO 3914; 1380pp + sequence listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been
CC   isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC   molecules have been determined. Primers for synthesising the full length
CC   cDNA are useful for clarifying the function of the protein encoded by
CC   the cDNA. The full length clones were obtained by construction of full
CC   length enriched cDNA libraries that were synthesised by the oligo-capping
CC   method. The primers enable the production of the full length cDNA easily
CC   without any special methods. The present sequence is a full length
CC   human cDNA of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ   Sequence 2989 BP; 795 A; 656 C; 698 G; 840 T; 0 other;


  Query Match             28.1%;  Score 999.4;  DB 22;  Length 2989;
  Best Local Similarity   99.9%;  Pred. No. 5.8e-262;
  Matches 1000;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2555 ttgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 2614
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 tggggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 257

Qy   2615 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 2674
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 317

Qy   2675 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 2734
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 377

Qy   2735 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 2794
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 437

Qy   2795 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 2854
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 497

Qy   2855 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 2914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 557

Qy   2915 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 2974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 617

Qy   2975 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 3034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 677

Qy   3035 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctaccca 3094
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    678 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctaccca 737

Qy   3095 accttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgcta 3154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    738 accttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgcta 797

Qy   3155 aacagttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcct 3214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    798 aacagttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcct 857

Qy   3215 ctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgca 3274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    858 ctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgca 917

Qy   3275 aacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttg 3334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    918 aacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttg 977

Qy   3335 aggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttacct 3394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978 aggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttacct 1037

Qy   3395 tagaactggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcag 3454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 tagaactggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcag 1097

Qy   3455 agcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctct 3514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 agcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctct 1157
```

```
Qy   3515 gcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          ||||||||||||||||||||||||||||||||||||||||||
Db   1158 gcaggagagattgcaacaccatcccacactgtccaggcctt 1198


RESULT  10
AAS34785
ID   AAS34785 standard; cDNA; 1269 BP.
XX
AC   AAS34785;
XX
DT   04-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human neoplastic disease associated polypeptide #19.
XX
KW   Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW   hyperproliferative disorder; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155163-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01358.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
```

```
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465558/50.
DR   P-PSDB; AAU21586.
XX
PT   Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT   diagnose diseases or disorders associated with aberrant expression or
PT   activity of polypeptides, and for treating cancers, rheumatoid
PT   arthritis -
XX
PS   Claim 4; SEQ ID No 29; 687pp; English.
XX
CC   The present invention relates to the isolation of novel human neoplastic
CC   disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
CC   sequences encoding for these polypeptides. The sequences of the
CC   invention are useful in the diagnosis, treatment, prevention and/or
CC   prognosis of disorders involving neoplastic disease such as
CC   hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC   cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC   astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC   also be useful for treating other disorders such as neural disorders,
CC   immune system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC   and renal disorders. The polynucleotide sequences of the invention are
CC   also useful in gene therapy. AAS34767-AAS35050 represent cDNA sequences
CC   encoding for the novel human neoplastic disease associated polypeptides
CC   of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1269 BP; 362 A; 313 C; 277 G; 312 T; 5 other;

  Query Match           26.5%;  Score 943.2;  DB 22;  Length 1269;
  Best Local Similarity  98.7%;  Pred. No. 7.7e-247;
  Matches  967;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    2;

Qy   1908 aaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagac 1967
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 aagggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagac 351

Qy   1968 ccggaggcattcctggaggcagcagatattcctccgagtagccacccccgcagaaggcgtg 2027
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 ccggaggcattcctggaggcagcagatattcctccgagtagccacccccgcagaaggcgtg 411

Qy   2028 cgattcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctcc 2087
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    412 cgattcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctcc 471

Qy   2088 tttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaaaggac 2147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 tttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaaaggac 531

Qy   2148 atctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttag 2207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    532 atctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttag 591

Qy   2208 aatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcct 2267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    592 aatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcct 651

Qy   2268 gaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaa 2327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    652 gaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaa 711

Qy   2328 gatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggc 2387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    712 gatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggc 771

Qy   2388 tgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagca 2447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    772 tgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagca 831

Qy   2448 attccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaaga 2507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    832 attccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaaga 891

Qy   2508 actcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctt 2567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    892 actcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctt 951

Qy   2568 tcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacat 2627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    952 tcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacat 1011

Qy   2628 tttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgt 2687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1012 tttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgt 1071

Qy   2688 agcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttct 2747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1072 agcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttct 1131

Qy   2748 gatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatcca 2807
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   1132 gatgtttgacatggggctgcggaaacagtatcggccagac-tgattattttacagatcca 1190

Qy   2808 gatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctgga 2867
           |||||:||||||:||||:|||||||||||||||||||| | |||||| ||||||||||
Db   1191 gatgtwccagctytcgargttgcttcatgattaccacagana-ctctaccatcacctgga 1249

Qy   2868 ggagcacgagatcggcccca 2887
           || ||||| ||| | ||||
Db   1250 agaacacgaaatcnggccca 1269


RESULT  11
AAH04450
ID   AAH04450 standard; cDNA; 883 BP.
XX
AC   AAH04450;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:1285.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
```

```
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 1; SEQ ID 1285; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 883 BP; 245 A; 221 C; 219 G; 193 T; 5 other;


  Query Match             22.1%;  Score 785;  DB 22;  Length 883;
  Best Local Similarity   96.7%;  Pred. No. 9.9e-204;
  Matches  843;  Conservative    0;  Mismatches   23;  Indels    6;  Gaps    4;

Qy   1444 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 1503
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 60

Qy   1504 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 1563
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 120

Qy   1564 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 1623
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 180

Qy   1624 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaag 1683
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaag 240

Qy   1684 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 1743
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 300

Qy   1744 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 1803
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 360

Qy   1804 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 1863
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 420

Qy   1864 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 1923
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 480

Qy   1924 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 1983
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 540

Qy   1984 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 2043
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 600

Qy   2044 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 2103
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    601 tatgaagattattcagagctgggagagctt-ccccacgatctcctttagaaccagtttgt 659

Qy   2104 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 2163
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db    660 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctncga 719

Qy   2164 gagc-tgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaa 2222
             ||| ||||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db    720 aagcttgtggcaaaaggctattcttcaacagatactgntgcttaaaatggagaaggaaaa 779

Qy   2223 tcagaagctccaagcc---tctgaaaatgatttgctgaacaagcgcc-tgaagctcgatt 2278
            ||||||||| | |  |  |   |||  ||||   |||||||||||| ||||  ||||||
Db    780 tcagaagcttccaagccttttgaaaatggattgcttgaacaagcgccttgaacttcgatt 839

Qy   2279 atgaagaaattactccctgtcttaaagaagta 2310
            ||||||||||||||| | ||||||||||
Db    840 ntgaagaaattactcctggctttaaagaagta 871


RESULT  12
AAK92006
ID   AAK92006 standard; cDNA; 849 BP.
XX
AC   AAK92006;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human cDNA 5'-end sequence, SEQ ID NO: 466.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-0114089.
XX
PR   08-JUL-1999;   99JP-0194486.
PR   11-JAN-2000; 2000JP-0118774.
PR   02-MAY-2000; 2000JP-0183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
DR   WPI; 2001-524255/58.
XX
```

PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 2; SEQ ID NO 466; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is the nucleotide
CC sequence of the 5'-end of a cDNA provided in the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;

Query Match 16.2%; Score 574.2; DB 22; Length 849;
Best Local Similarity 98.2%; Pred. No. 3.5e-146;
Matches 642; Conservative 0; Mismatches 6; Indels 6; Gaps 6;

```
Qy   2555 ttgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 2614
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 tggggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 257

Qy   2615 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 2674
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 317

Qy   2675 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 2734
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 377

Qy   2735 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 2794
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 437

Qy   2795 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 2854
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 497

Qy   2855 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 2914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 557

Qy   2915 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 2974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 617

Qy   2975 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 3034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 677

Qy   3035 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctaccca 3094
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    678 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgct-ccca 736

Qy   3095 accttggcttggtacagatggaaaagaccatcaatcaggta-tttgaaatggacatc-gc 3152
          |||||||||||| |||||||||||||||||||||||| ||| |||| |||||||||| |
Db    737 accttggcttgg-acagatggaaaagaccatcaatcangtattttgnaatggacatcggt 795

Qy   3153 taaacagttacaagcttatg-aagttgagtacca-cgtccttcaagaagaactt 3204
          |||||||||||||||||| |||||||| ||| ||||||| ||||||||||
Db    796 taaacagttacaagcttatgaaagttgagtnccaccgtcctttaagaagaactt 849
```

RESULT 13
AAK93666
ID AAK93666 standard; cDNA; 849 BP.
XX
AC AAK93666;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA clone representative sequence, SEQ ID NO: 2126.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Example 11; SEQ ID NO 2126; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used as the
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;

Query Match 16.2%; Score 574.2; DB 22; Length 849;
Best Local Similarity 98.2%; Pred. No. 3.5e-146;
Matches 642; Conservative 0; Mismatches 6; Indels 6; Gaps 6;

```
Qy   2555 ttgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 2614
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 tggggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 257

Qy   2615 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 2674
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 317

Qy   2675 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 2734
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 377

Qy   2735 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 2794
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 437

Qy   2795 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 2854
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 497
```

```
Qy   2855 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 2914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 557

Qy   2915 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 2974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 617

Qy   2975 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 3034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 677

Qy   3035 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctaccca 3094
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    678 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgct-ccca 736

Qy   3095 accttggcttggtacagatggaaaagaccatcaatcaggta-tttgaaatggacatc-gc 3152
          |||||||||||| |||||||||||||||||||| ||| |||| |||||||||| |
Db    737 accttggcttgg-acagatggaaaagaccatcaatcangtattttgnaatggacatcggt 795

Qy   3153 taaacagttacaagcttatg-aagttgagtacca-cgtccttcaagaagaactt 3204
          |||||||||||||||||| ||||||||| ||| ||||||| ||||||||||
Db    796 taaacagttacaagcttatgaaagttgagtnccaccgtcctttaagaagaactt 849
```

RESULT 14
AAZ50904
ID AAZ50904 standard; DNA; 17590 BP.
XX
AC AAZ50904;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human TBC-1 partial genomic DNA comprising 5' end sequence.
XX
KW TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;
KW Single nucleotide polymorphism; tissue differentiation; prostate cancer;
KW linkage analysis; genetic map; detection; diagnosis; genotyping;
KW transgenic animal; screening; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	misc_signal	1..2000
FT		/*tag= a
FT		/note= "5' Regulatory region"
FT	exon	2001..2077
FT		/*tag= b
FT		/number= 1
FT	intron	2078..12739
FT		/*tag= c
FT		/number= 1
FT	misc_feature	9391..9845
FT		/*tag= d
FT		/note= "Amplicon 99-430"
FT	primer_bind	9391..9408
FT		/*tag= e
FT		/bound_moiety= "Primer B1"
FT		/note= "Amplification of amplicon 99-430"
FT	primer_bind	9475..9493
FT		/*tag= f
FT		/bound_moiety= "Primer D1"
FT		/note= "Microsequencing of marker 99-430-352"
FT	misc_binding	9482..9506
FT		/*tag= g
FT		/bound_moiety= "Probe P1"
FT		/note= "Detection of Biallelic marker 99-430-352"
FT	misc_feature	9494
FT		/*tag= h
FT		/note= "Ambiguity base 'R' corresponds to 'A' in allele-1
FT		and 'G' in allele-2 of biallelic marker 99-430-352"
FT	primer_bind	complement (9495..9513)
FT		/*tag= i
FT		/bound_moiety= "Primer E1"
FT		/note= "Microsequencing of marker 99-430-352"
FT	primer_bind	complement (9828..9845)
FT		/*tag= j
FT		/bound_moiety= "Primer C1"
FT		/note= "Amplification of amplicon 99-430"
FT	exon	12292..12373
FT		/*tag= k
FT		/number= 1 bis
FT	intron	12374..12739
FT		/*tag= l
FT		/number= 1 bis
FT	exon	12740..13249
FT		/*tag= m
FT		/number= 2
FT	intron	13250..17590
FT		/*tag= n
FT		/number= 2

XX
PN WO200008209-A2.
XX
PD 17-FEB-2000.
XX
PF 06-AUG-1999; 99WO-IB01444.
XX
PR 07-AUG-1998; 98US-0095653.
XX
PA (GEST) GENSET.
XX
PI Blumenfeld M, Bougueleret L, Chumakov I;
XX
DR WPI; 2000-205736/18.
XX
PT New isolated human TBC-1 nucleic acids, useful for developing products
PT for the diagnosis and treatment of disorders involving cell
PT proliferation, particularly prostate cancer
XX
PS Claim 1; Page 93-100; 166pp; English.
XX
CC The present sequence is the partial genomic DNA of human TBC-1 gene,
CC comprising the 5' regulatory region, exons 1, 1bis and 2. TBC-1 gene is
CC mapped to a candidate region of prostate cancer on chromosome 4. Single
CC nucleotide polymorphism (SNP) is located within the biallelic marker
CC region 99-430-352, localised in intron 1 of TBC-1 genomic DNA.
CC TBC-1 gene is involved in the regulation of cell cycle and tissue
CC differentiation in mammals. An alteration of TBC-1 sequence may be
CC associated with a pathological condition, resulting in abnormal cell
CC proliferation leading to cancer, e.g. prostate cancer. The biallelic
CC markers can be used for generation of genetic maps, linkage analysis and
CC association studies. TBC-1 sequence can be used for detection,
CC diagnosis, genotyping, production of transgenic animals and screening
CC of compounds for use in therapy.
XX
SQ Sequence 17590 BP; 4760 A; 3776 C; 4104 G; 4919 T; 31 other;

Query Match 11.9%; Score 422.6; DB 21; Length 17590;
Best Local Similarity 95.8%; Pred. No. 5.4e-104;
Matches 434; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy      1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  12833 atggaaccaataacattcacagcaaggaaacatctgctttctaacgaggtctcggtggat 12892

Qy     61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12893 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 12952

Qy    121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12953 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 13012
```

```
Qy     181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13013 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 13072

Qy     241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13073 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 13132

Qy     301 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13133 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 13192

Qy     361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13193 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagta 13252

Qy     421 cctgagatcatcagctccatccgtcaggcgggg 453
             ||||||    | |   |   | ||| | ||
Db   13253 agtgagatggagatccaaaagactaaggtgtgg 13285
```

RESULT 15
AAZ15229
ID AAZ15229 standard; cDNA; 696 BP.
XX
AC AAZ15229;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:2698.
XX
KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US01619.
XX
PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
DR WPI; 1999-494092/41.
XX
PT Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX
PS Claim 1; Page 1311; 2479pp; English.
XX
CC The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.
XX
SQ Sequence 696 BP; 189 A; 157 C; 178 G; 142 T; 30 other;

Query Match 10.2%; Score 362; DB 20; Length 696;
Best Local Similarity 100.0%; Pred. No. 2.8e-88;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    3194 aagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagaga 3253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59 aagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagaga 118

Qy    3254 aaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaa 3313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 aaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaa 178

Qy    3314 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 3373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     179 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 238

Qy    3374 tgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggagg 3433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     239 tgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggagg 298

Qy    3434 agctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagc 3493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 agctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagc 358

Qy    3494 ccacgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcc 3553
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     359 ccacgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcc 418

Qy    3554 tt 3555
           ||
Db     419 tt 420
```

Search completed: September 9, 2002, 19:20:22
Job time: 17600 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 19:16:17 ; Search time 8477.27 Seconds
(without alignments)
8775.683 Million cell updates/sec

Title: US-09-762-311-4_COPY_176_3730
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2751	77.4	3023	9	BC014529	BC014529 Homo sapi
2	2510.4	70.6	4039	6	I86429	I86429 Sequence 1
3	2510.4	70.6	4039	10	MMU33005	U33005 Mus musculu
4	2340	65.8	2576	9	AB029031	AB029031 Homo sapi
5	2112	59.4	2362	9	AK027355	AK027355 Homo sapi
6	1359.2	38.2	1995	4	BTA17923	Y17923 Bos taurus
7	1179.6	33.2	3051	10	BC004675	BC004675 Mus muscu
8	998.4	28.1	1788	9	AK057182	AK057182 Homo sapi
9	682.2	19.2	5922	9	AB011175	AB011175 Homo sapi
c 10	465.8	13.1	189811	30	AC009595	Ac009595 Homo sapi
11	465.8	13.1	195108	9	AC021106	AC021106 Homo sapi
c 12	465.8	13.1	208318	2	AC108933	AC108933 Homo sapi
c 13	424.2	11.9	110000	2	AL390202_04	Continuation (5 of
14	304.6	8.6	3214	3	DMU50542	U50542 Drosophila
15	304.6	8.6	4746	3	DME17919	Y17919 Drosophila
16	284.4	8.0	189811	30	AC009595	Ac009595 Homo sapi
17	284.4	8.0	208318	2	AC108933	AC108933 Homo sapi
18	252.4	7.1	163928	2	AC021335	AC021335 Homo sapi
19	252.4	7.1	180918	9	AC098680	AC098680 Homo sapi
20	245.6	6.9	17019	2	AC019886	AC019886 Drosophil
21	245.6	6.9	161117	3	AC008311	AC008311 Drosophil
22	245.6	6.9	168448	3	AC009346	AC009346 Drosophil
23	245.6	6.9	297832	3	AE003602	AE003602 Drosophil
c 24	180	5.1	163928	2	AC021335	AC021335 Homo sapi
25	141.8	4.0	3761	9	BC014111	BC014111 Homo sapi
c 26	141	4.0	221717	2	AC074357	AC074357 Mus muscu
c 27	139.2	3.9	158310	9	AL162571	AL162571 Human DNA
28	139.2	3.9	183577	2	AC011164	AC011164 Homo sapi
29	132.8	3.7	3691	6	AX078376	AX078376 Sequence
30	132.8	3.7	4560	9	HSA011679	AJ011679 Homo sapi
31	123.8	3.5	154	6	AX261661	AX261661 Sequence
32	119.4	3.4	14119	2	AC013898	AC013898 Drosophil
c 33	119.4	3.4	164413	3	AC023718	AC023718 Drosophil
c 34	119.4	3.4	303648	3	AE003486	AE003486 Drosophil
35	114.6	3.2	2013	9	AF042345	AF042345 Homo sapi
36	114.6	3.2	7383	9	AF008915	AF008915 Homo sapi
c 37	101.8	2.9	30027	3	CELT24D11	U49940 Caenorhabdi
c 38	98	2.8	20553	3	CBRG26M15	AC084540 Caenorhab
c 39	95.6	2.7	73803	8	NCB11B22	AL356834 Neurospor
40	90	2.5	5870	10	MMU53586	U53586 Mus musculu
c 41	84.2	2.4	441	11	G47994	G47994 Z27058_1 Ze
42	80	2.3	27102	3	U41540	U41540 Caenorhabdi
c 43	79.6	2.2	110032	2	AC014916	AC014916 Drosophil
44	79.6	2.2	159660	3	AC010040	AC010040 Drosophil
45	79.6	2.2	282267	3	AE003555	AE003555 Drosophil

ALIGNMENTS

RESULT 1
BC014529
LOCUS BC014529 3023 bp mRNA linear PRI 19-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:3843156, mRNA.
ACCESSION BC014529
VERSION BC014529.1 GI:17939551
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3023)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 14 Row: c Column: 16
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

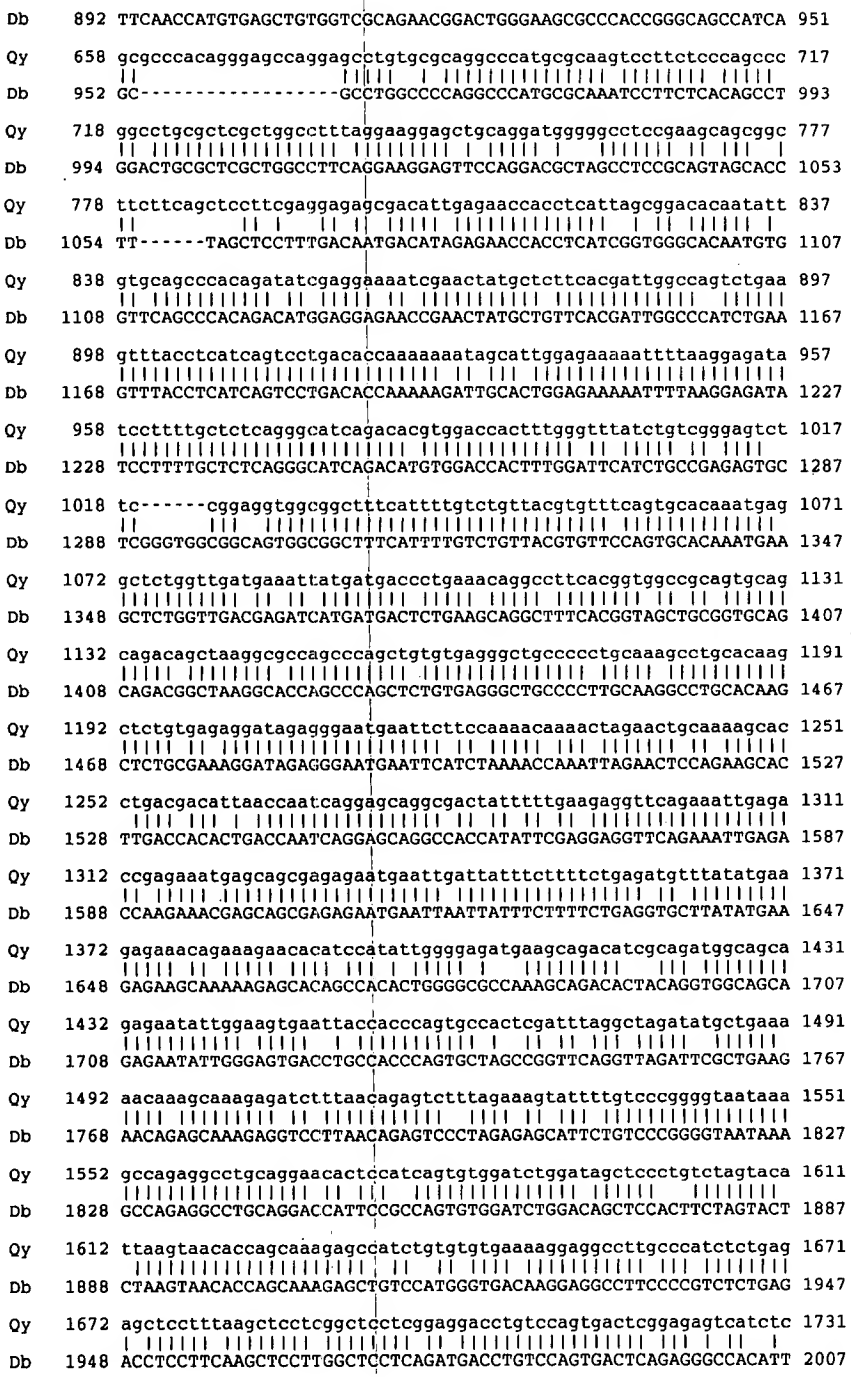
FEATURES Location/Qualifiers
source 1. .3023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3843156"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
BASE COUNT 883 a 712 c 740 g 688 t
ORIGIN

Query Match 77.4%; Score 2751; DB 9; Length 3023;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    805 attgagaaccacctcattagcggacacaatattgtgcagcccacagatatcgaggaaaat 864
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 ATTGAGAACCACCTCATTAGCGGACACAATATTGTGCAGCCCACAGATATCGAGGAAAAT 69

Qy    865 cgaactatgctcttcacgattggccagtctgaagtttacctcatcagtcctgacaccaaa 924
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 CGAACTATGCTCTTCACGATTGGCCAGTCTGAAGTTTACCTCATCAGTCCTGACACCAAA 129

Qy    925 aaaatagcattggagaaaaatttttaaggagatatccttttgctctcagggcatcagacac 984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 AAAATAGCATTGGAGAAAAATTTTAAGGAGATATCCTTTTGCTCTCAGGGCATCAGACAC 189

Qy    985 gtggaccactttgggtttatctgtcgggagtcttccggaggtggcggctttcattttgtc 1044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 GTGGACCACTTTGGGTTTATCTGTCGGGAGTCTTCCGGAGGTGGCGGCTTTCATTTTGTC 249

Qy   1045 tgttacgtgtttcagtgcacaaatgaggctctggttgatgaaattatgatgaccctgaaa 1104
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 TGTTACGTGTTTCAGTGCACAAATGAGGCTCTGGTTGATGAAATTATGATGACCCTGAAA 309

Qy   1105 caggccttcacggtggccgcagtgcagcagacagctaaggcgccagcccagctgtgtgag 1164
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 CAGGCCTTCACGGTGGCCGCAGTGCAGCAGACAGCTAAGGCGCCAGCCCAGCTGTGTGAG 369

Qy   1165 ggctgccccctgcaaagcctgcacaagctctgtgagaggatagagggaatgaattcttcc 1224
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 GGCTGCCCCCTGCAAAGCCTGCACAAGCTCTGTGAGAGGATAGAGGGAATGAATTCTTCC 429

Qy   1225 aaaacaaaactagaactgcaaaagcacctgacgacattaaccaatcaggagcaggcgact 1284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 AAAACAAAACTAGAACTGCAAAAGCACCTGACGACATTAACCAATCAGGAGCAGGCGACT 489

Qy   1285 atttttgaagaggttcagaaattgagaccgagaaatgagcagcgagagaatgaattgatt 1344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 ATTTTTGAAGAGGTTCAGAAATTGAGACCGAGAAATGAGCAGCGAGAGAATGAATTGATT 549

Qy   1345 atttcttttctgagatgtttatatgaagagaaacagaaagaacacatccatattggggag 1404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    550 ATTTCTTTTCTGAGATGTTTATATGAAGAGAAACAGAAAGAACACATCCATATTGGGGAG 609

Qy   1405 atgaagcagacatcgcagatggcagcagagaatattggaagtgaattaccacccagtgcc 1464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 ATGAAGCAGACATCGCAGATGGCAGCAGAGAATATTGGAAGTGAATTACCACCCAGTGCC 669

Qy   1465 actcgatttaggctagatatgctgaaaaacaaagcaaagagatctttaacagagtcttta 1524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    670 ACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAGAGATCTTTAACAGAGTCTTTA 729

Qy   1525 gaaagtattttgtcccggggtaataaagccagaggcctgcaggaacactccatcagtgtg 1584
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    730 GAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTGCAGGAACACTCCATCAGTGTG 789

Qy   1585 gatctggatagctccctgtctagtacattaagtaacaccagcaaagagccatctgtgtgt 1644
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    790 GATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCATCTGTGTGT 849

Qy   1645 gaaaaggaggccttgcccatctctgagagctcctttaagctcctcggctcctcggaggac 1704
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    850 GAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCCTCGGAGGAC 909

Qy   1705 ctgtccagtgactcggagagtcatctcccagaagagccagctccgctgtcgccccagcag 1764
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    910 CTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCGCCCCAGCAG 969

Qy   1765 gccttcaggaggcgagcaaacaccctgagtcacttccccatcgaatgccaggaacctcca 1824
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    970 GCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAGGAACCTCCA 1029

Qy   1825 caacctgcccgggggtccccgggggtttcgcaaaggaaacttatgaggtatcactcagtg 1884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1030 CAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTATCACTCAGTG 1089

Qy   1885 agcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatcttggtgat 1944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1090 AGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTTGGTGAT 1149

Qy   1945 tctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatattcctccga 2004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1150 TCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATTCCTCCGA 1209

Qy   2005 gtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattcagagctg 2064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1210 GTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTCAGAGCTG 1269

Qy   2065 ggagagcttcccccacgatctccttttagaaccagtttgtgaagatgggccctttggcccc 2124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1270 GGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCCTTTGGCCCC 1329

Qy   2125 ccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaaggctatt 2184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1330 CCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAAGGCTATT 1389

Qy   2185 cttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaagcctctgaa 2244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1390 CTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGCCTCTGAA 1449

Qy   2245 aatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccctgtcttaaa 2304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1450 AATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCCTGTCTTAAA 1509

Qy   2305 gaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaaattaagttt 2364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1510 GAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAAATTAAGTTT 1569

Qy   2365 gacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaa 2424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1570 GACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCATCACCGAGGTGAA 1629

Qy   2425 atctggaaatttctagctgagcaattccaccttaaacaccagtttcccagcaaacagcag 2484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1630 ATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGCAAACAGCAG 1689

Qy   2485 ccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcagcatgcgatt 2544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1690 CCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAGCATGCGATT 1749
```

```
Qy   2545 cttattgaccttgggcgaaccttcctacacacccatacttctctgcccagcttggagca 2604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1750 CTTATTGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAGCTTGGAGCA 1809

Qy   2605 ggacagctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtggga 2664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1810 GGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAGACCAGGAAGTGGGA 1869

Qy   2665 tattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagag 2724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1870 TATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTTTGCTTCTTCATATGAGTGAGGAAGAG 1929

Qy   2725 gcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggcca 2784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1930 GCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAAACAGTATCGGCCA 1989

Qy   2785 gacatgattattttacagatccagatgtaccagctctcgaggttgcttcatgattaccac 2844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1990 GACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCTTCATGATTACCAC 2049

Qy   2845 agagacctctacaatcacctggaggagcacgagatcggccccagcctctacgctgccccc 2904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2050 AGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCTCTACGCTGCCCCC 2109

Qy   2905 tggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgat 2964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2110 TGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCAGAGTCTTTGAT 2169

Qy   2965 atgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagc 3024
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2170 ATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTCTGTTGGGAAGC 2229

Qy   3025 cataagcccttgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagc 3084
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2230 CATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACTTTATAAAAAGC 2289

Qy   3085 acgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatg 3144
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2290 ACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGGTATTTGAAATG 2349

Qy   3145 gacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaagaagaactt 3204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2350 GACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTCAAGAAGAACTT 2409

Qy   3205 atcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagc 3264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2410 ATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGAAAACCAACAGC 2469

Qy   3265 agcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatc 3324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2470 AGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAAATGGTAGGATC 2529

Qy   3325 caaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggcc 3384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2530 CAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGCTGAAGCAGGCC 2589

Qy   3385 atgcttaccttagaactggagcggtcggccctgctgcagacggtggaggagctgcggcgg 3444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2590 ATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGGAGCTGCGGCGG 2649

Qy   3445 cggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgac 3504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2650 CGGAGCGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGCCCACGGGCGAC 2709

Qy   3505 tgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2710 TGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCCTT 2760


RESULT   2
I86429
LOCUS       I86429                 4039 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 1 from patent US 5700927.
ACCESSION   I86429
VERSION     I86429.1  GI:3206147
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4039)
  AUTHORS   Zon,L. and Richardson,P.
  TITLE     Tbc1 gene and uses thereof
  JOURNAL   Patent: US 5700927-A 1 23-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .4039
                     /organism="unknown"
BASE COUNT     1042 a   1096 c   1060 g    841 t
ORIGIN


  Query Match             70.6%;  Score 2510.4;  DB 6;  Length 4039;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches 2949;  Conservative    0;  Mismatches  536;  Indels   37;  Gaps    8;

Qy      1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59
          |||||  |||| |||||||||||| ||||| |||| | || |||||||| ||||| |||||
Db    292 ATGGAGGCAATCACATTCACAGCGAGGAAGCATCCGTTTCCCTAACGAAGTCTCCGTGGA 351

Qy     60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
           || ||| |||||||||  || |||| | || |||||||| || ||||| |||||||||
Db    352 CTTCGGCTTGCAGCTGGTTGGGTCTCCTTACCCGTGCATTCTCTCACCACTATGCCCATG 411

Qy    118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
          ||||| ||||| || || ||||| ||||||||||  | |||| | | | |||||| |||
Db    412 CTGCCGTGGGTGGTAGCCGAGGTACGAAGACTCAGTGGCCAGTGCTCTAAAAAGGAGCCT 471

Qy    178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
            || |||||||||| ||||| ||||| ||||| |||||| | || || |||||| |
Db    472 AGGACAAAGCAAGTCCGCCTTTGGGTTTCGCCCTCCGGACTGCGGTGCGAGCCTGACCTG 531

Qy    238 gggagaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcag 297
          | || |||||||  ||||| || || |||| ||||||||||||||||||||||||||||
Db    532 GAGAAAAGTCAACCATGGGACCCGCTCATCTGTTCCAGCATCTTTGAGTGCAAGCCTCAG 591

Qy    298 cgtgttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
          ||||| |||||||||||||||||| ||||| || |||||||||| || ||||||
Db    592 CGTGTGCACAAACTGATTCACAACAGTCACGACCCGAGCTACTTTGCTTGCCTCATTAAG 651

Qy    358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
          || || |||| ||| |||||||  |||||||||||| ||||| || ||||||||||||
Db    652 GAGGATGCTGCGCACAGGCAGAGCCTCTGCTATGTGTTTAAAGCAGACGATCAAACAAAA 711

Qy    418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
          |||||||||||||||||||||||| ||||| |||||||| ||||| ||||| ||||||
Db    712 GTGCCTGAGATCATCAGCTCCATCCGGCAGGCCGGGAAGATTGCCCGCCAGGAAGAGCTG 771

Qy    478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
          |  ||||| |||||||||||||| || ||  |||| |||||||||||||||||||| |||
Db    772 CGTTGCCCCTCCGAGTTCGACGATACCTTCGCCAAAAAGTTCGAGGTGCTCTTCTGTGGC 831

Qy    538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
          || ||||| ||||| |||||||||| || || || |||| |||||||| |||||||||
Db    832 CGGGTGACTGTGGCTCACAAGAAGGCCCCACCCGCACTGATTGACGAGTGTATCGAGAAG 891

Qy    598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
          ||||| || || ||| |  | ||  |  | ||  |       |||| |  ||  |  |
Db    892 TTCAACCATGTGAGCTGTGGTCGCAGAACGGACTGGGAAGCGCCCACCGGGCAGCCATCA 951

Qy    658 gcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagccc 717
          ||                    |||||  | |||||||||||||| ||||||| |||||
Db    952 GC------------------GCCTGGCCCCAGGCCCATGCGCAAATCCTTCTCACAGCCT 993

Qy    718 ggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggc 777
          || |||||||||||||||| |||||||| | ||||| |   |||||| || |||  |
Db    994 GGACTGCGCTCGCTGGCCTTCAGGAAGGAGTTCCAGGACGCTAGCCTCCGCAGTAGCACC 1053
```

```
Db   3207 GTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGATCAGAGGT 3266

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          ||||||||||| |||||||||| ||||| ||||||||||||||||||| ||||||||
Db   3267 CATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTACAGCATGA 3326

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
           ||||| |||||||| || ||||| ||||| | ||| || |||||||| ||| |||| ||
Db   3327 GAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGCCTGGTGCA 3386

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          |||||| |||||||||| |||||| ||||| |||||||||||| || ||| | || || ||
Db   3387 GATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTCCAGGCCTA 3446

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||| || |||||||||||| | || || || ||||| || ||||| ||||||||||||||
Db   3447 TGAGGTCGAGTACCACGTCGTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTCAGTGACAA 3506

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          |||||||||||| ||||| ||||||||||||||||  || ||||||||||||||||||||
Db   3507 CCAAAGAATGGAGAAATTGGAGAAAACCAACAGCACGTTGCGCAAACAGAACCTTGACCT 3566

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
          ||| || ||||||||||||||||||| |||||||||||||||||| |||||  | |||||
Db   3567 CCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACGGTAGAGAA 3626

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
           || || | ||| ||||| |||||||||||     ||| ||| | ||  |||||||
Db   3627 ACTTCTTACCAGCGAGAGTAAGCTGAAGCAGCGTGCGCTGACCCTGGAGGTGGAGCG--T 3684

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
          |||||||||||||| |||||||||||||||| |||  |||||   |||||||   | ||
Db   3685 CGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGCACTCCAGA 3744

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
          ||| || ||||| ||||  |||||||| ||||| |||| |||
Db   3745 GCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 3786


RESULT   3
MMU33005
LOCUS       MMU33005                4039 bp    mRNA    linear   ROD 12-JAN-1996
DEFINITION  Mus musculus tbc1 mRNA, complete cds.
ACCESSION   U33005
VERSION     U33005.1  GI:988220
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4039)
  AUTHORS   Richardson,P.M. and Zon,L.I.
  TITLE     Molecular cloning of a cDNA with a novel domain present in the
            tre-2 oncogene and the yeast cell cycle regulators BUB2 and cdc16
  JOURNAL   Oncogene 11 (6), 1139-1148 (1995)
  MEDLINE   96032578
   PUBMED   7566974
REFERENCE   2  (bases 1 to 4039)
  AUTHORS   Richardson,P.M. and Zon,L.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-1995) Leonard I. Zon, Hematology, Childrens
            Hospital of Boston, HHMI, 300 Longwood Ave., Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1. .4039
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_type="mast cells"
     gene            1. .4039
                     /gene="tbc1"
     CDS             403. .3828
                     /gene="tbc1"
                     /note="ORF1"
                     /codon_start=1
                     /product="Tbc1"
                     /protein_id="AAA85223.1"
                     /db_xref="GI:988221"
                     /translation="MPMLPWVVAEVRRLSGQCSKKEPRTKQVRLWVSPSGLRCEPDLE
                     KSQPWDPLICSSIFECKPQRVHKLIHNSHDPSYFACLIKEDAAHRQSLCYVFKADDQT
                     KVPEIISSIRQAGKIARQEELRCPSEFDDTFAKKFEVLFCGRVTVAHKKAPPALIDEC
                     IEKFNHVSCGRRTDWEAPTGQPSAPGPRPMRKSFSQPGLRSLAFRKEFQDASLRSSTF
                     SSFDNDIENHLIGGHNVVQPTDMEENRTMLFTIGPSEVYLISPDTKKIALEKNFKEIS
                     FCSQGIRHVDHFGFICRECSGGGSGGFHFVCYVFQCTNEALVDEIMMTLKQAFTVAAV
                     QQTAKAPAQLCEGCPLQGLHKLCERIEGMNSSKTKLELQKHLTTLTNQEQATIFEEVQ
                     KLRPRNEQRENELIISFLRCLYEEKQKEHSHTGAPKQTLQVAAENIGSDLPPSASRFR
                     LDSLKNRAKRSLTESLESILSRGNKARGLQDHSASVDLDSSTSSTLSNTSKELSMGDK
                     EAFPVSETSFKLLGSSDDLSSDSEGHIAEESALLSPQQAFRRRANTLSHFPVECPAPP
                     EPAQSSPGVSQRKLMRYHSVSTETPHERKDFESKANHLGDTDGTPVKTRRHSWRQQIF
                     LRVATPQKACDSPSRYEDYSELGELPPRSPLEPVCEDGPFGQYRKKRGRRHASFESCG
                     KRPSCSRSCLVRMEKENQKLQASENDLLNKRLKLDYEEITPCLKEVTTVWEKMLSTPG
                     RSKIKFDMEKVHSAVGQGVPRHHRGEIWKFLAEQFHLKHPFPSKQQPKDVPYKELLKK
                     LTSQQHAILIDLGRTFPTHPYFSAQLGAGQLSLYNILKAYSLLDQEVGYCQGLSFVAG
                     ILLLHMSEEEAFKMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHDYHRDLYNHLE
                     EHETGPPTYAAPWFLTVFASQFPLGFVARVFDMIFLQGSEVIFKVALSLLGSHKPLIL
                     QHENLETIVDFIKNTLPNLGLVQMEKTISQVFEMDIAKQLQAYEVEYHVVQEELIESS
                     PLSDNQRMEKLEKTNSTLRKQNLDLLEQLQVANARIQSLEATVEKLLTSESKLKQRAL
                     TLEVERRPAADGGGAAEAKRPAQHSRARLHPAGAHRRLTAARRDCAPLTLSKP"
BASE COUNT     1042 a    1096 c    1060 g     841 t
ORIGIN


  Query Match             70.6%;  Score 2510.4;  DB 10;  Length 4039;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches 2949;  Conservative    0;  Mismatches  536;  Indels   37;  Gaps    8;

Qy       1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59
           ||||| |||| |||||||||||| ||||| |||| | || |||||||| ||||| |||||
Db     292 ATGGAGGCAATCACATTCACAGCGAGGAAGCATCCGTTTCCCTAACGAAGTCTCCGTGGA 351

Qy      60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
            || ||| |||||||||||  || |||| | || |||||||| || ||||| |||||||||
Db     352 CTTCGGCTTGCAGCTGGTTGGGTCTCCTTACCCGTGCATTCTCTCACCACTATGCCCATG 411

Qy     118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
           ||||| ||||| || || ||||| ||||||||||||  | |||| | | | |||||| |||
Db     412 CTGCCGTGGGTGGTAGCCGAGGTACGAAGACTCAGTGGCCAGTGCTCTAAAAAGGAGCCT 471

Qy     178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
             || |||||||||||| ||||| ||||| ||||| |||||| | || || ||||| |
Db     472 AGGACAAAGCAAGTCCGCCTTTGGGTTTCGCCCTCCGGACTGCGGTGCGAGCCTGACCTG 531

Qy     238 gggagaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcag 297
           | || |||||||  ||||| || || |||| ||||||||||||||||||||||||||||||
Db     532 GAGAAAAGTCAACCATGGGACCCGCTCATCTGTTCCAGCATCTTTGAGTGCAAGCCTCAG 591

Qy     298 cgtgttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
           ||||| |||||||||||||||||||||| ||||| || |||||||||| || ||||||
Db     592 CGTGTGCACAAACTGATTCACAACAGTCACGACCCGAGCTACTTTGCTTGCCTCATTAAG 651

Qy     358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
           || || ||||  ||| |||||||  |||||||||||| ||||| || ||||||||||||
Db     652 GAGGATGCTGCGCACAGGCAGAGCCTCTGCTATGTGTTTAAAGCAGACGATCAAACAAAA 711

Qy     418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
           |||||||||||||||||||||||||||| ||||| |||||||| ||||| |||||| ||||||
Db     712 GTGCCTGAGATCATCAGCTCCATCCGGCAGGCCGGGAAGATTGCCCGCCAGGAAGAGCTG 771

Qy     478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
           |  ||||| |||||||||||||| || ||  |||| |||||||||||||||||||| |||
Db     772 CGTTGCCCCTCCGAGTTCGACGATACCTTCGCCAAAAAGTTCGAGGTGCTCTTCTGTGGC 831

Qy     538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
           || ||||| ||||| |||||||||| || || || ||||| |||||||| ||||||||
Db     832 CGGGTGACTGTGGCTCACAAGAAGGCCCCACCCGCACTGATTGACGAGTGTATCGAGAAG 891

Qy     598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
           ||||| || || ||| |  | ||  |  | ||  |       |||| |  ||  |   |
```


Db 892 TTCAACCATGTGAGCTGTGGTCGCAGAACGGACTGGGAAGCGCCCACCGGGCAGCCATCA 951
Qy 658 gcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagccc 717
Db 952 GC----------GCCTGGCCCCAGGCCCATGCGCAAATCCTTCTCACAGCCT 993
Qy 718 ggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggc 777
Db 994 GGACTGCGCTCGCTGGCCTTCAGGAAGGAGTTCCAGGACGCTAGCCTCCGCAGTAGCACC 1053
Qy 778 ttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatatt 837
Db 1054 TT------TAGCTCCTTTGACAATGACATAGAGAACCACCTCATCGGTGGGCACAATGTG 1107
Qy 838 gtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaa 897
Db 1108 GTTCAGCCCACAGACATGGAGGAGAACCGAACTATGCTGTTCACGATTGGCCCATCTGAA 1167
Qy 898 gtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttaaggagata 957
Db 1168 GTTTACCTCATCAGTCCTGACACCAAAAAGATTGCACTGGAGAAAAATTTTAAGGAGATA 1227
Qy 958 tccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtct 1017
Db 1228 TCCTTTTGCTCTCAGGGCATCAGACATGTGGACCACTTTGGATTCATCTGCCGAGAGTGC 1287
Qy 1018 tc------cggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
Db 1288 TCGGGTGGCGGCAGTGGCGGCTTTCATTTTGTCTGTTACGTGTTCCAGTGCACAAATGAA 1347
Qy 1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
Db 1348 GCTCTGGTTGACGAGATCATGATGACTCTGAAGCAGGCTTTCACGGTAGCTGCGGTGCAG 1407
Qy 1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
Db 1408 CAGACGGCTAAGGCACCAGCCCAGCTCTGTGAGGGCTGCCCCTTGCAAGGCCTGCACAAG 1467
Qy 1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
Db 1468 CTCTGCGAAAGGATAGAGGGAATGAATTCATCTAAAACCAAATTAGAACTCCAGAAGCAC 1527
Qy 1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
Db 1528 TTGACCACACTGACCAATCAGGAGCAGGCCACCATATTCGAGGAGGTTCAGAAATTGAGA 1587
Qy 1312 ccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaa 1371
Db 1588 CCAAGAAACGAGCAGCGAGAGAATGAATTAATTATTTCTTTTCTGAGGTGCTTATATGAA 1647
Qy 1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
Db 1648 GAGAAGCAAAAAGAGCACAGCCACACTGGGGCGCCAAAGCAGACACTACAGGTGGCAGCA 1707
Qy 1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
Db 1708 GAGAATATTGGGAGTGACCTGCCACCCAGTGCTAGCCGGTTCAGGTTAGATTCGCTGAAG 1767
Qy 1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccgggggtaataaa 1551
Db 1768 AACAGAGCAAAGAGGTCCTTAACAGAGTCCCTAGAGAGCATTCTGTCCCGGGGTAATAAA 1827
Qy 1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
Db 1828 GCCAGAGGCCTGCAGGACCATTCCGCCAGTGTGGATCTGGACAGCTCCACTTCTAGTACT 1887
Qy 1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
Db 1888 CTAAGTAACACCAGCAAAGAGCTGTCCATGGGTGACAAGGAGGCCTTCCCCGTCTCTGAG 1947
Qy 1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
Db 1948 ACCTCCTTCAAGCTCCTTGGCTCCTCAGATGACCTGTCCAGTGACTCAGAGGGCCACATT 2007
Qy 1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
Db 2008 GCAGAAGAGTCTGCCCTGTTGTCACCCCAGCAGGCGTTCAGAAGGAGAGCCAACACCCTG 2067
Qy 1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
Db 2068 AGTCATTTCCCAGTAGAGTGCCCTGCGCCTCCAGAACCTGCCCAGAGCTCTCCAGGGGTC 2127
Qy 1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
Db 2128 TCTCAAAGGAAACTCATGCGGTACCACTCCGTGAGCACAGAGACGCCTCATGAACGCAAG 2187
Qy 1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
Db 2188 GACTTTGAATCCAAAGCAAACCACCTGGGTGACACAGATGGGACCCCCGTGAAGACCCGG 2247
Qy 1972 aggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgat 2031
Db 2248 CGGCACTCGTGGAGACAGCAGATATTCCTTCGAGTGGCCACTCCACAGAAGGCTTGTGAC 2307
Qy 2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
Db 2308 TCCCCGAGCAGATATGAAGATTATTCCGAGCTGGGAGAGCTCCCTCCACGCTCCCCTTTA 2367
Qy 2092 gaaccagtttgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatct 2151
Db 2368 GAACCGGTGTGTGAGGACGGCCCATTTGG-CCAGTACAGGAAGAAAAGAGGAAGACGTCA 2426
Qy 2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg-ctgcttagaat 2210
Db 2427 CGCGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTGCAGCAGATCCTGCCTCGTCAGGAT 2486
Qy 2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
Db 2487 GGAGAAGGAGAATCAGAAGCTACAAGCCTCTGAAAACGATTTGCTGAACAAACGCCTCAA 2546
Qy 2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
Db 2547 GCTTGACTATGAAGAAATCACTCCGTGTCTTAAAGAAGTCACTACAGTGTGGGAAAAGAT 2606
Qy 2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
Db 2607 GCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAAGTGCACTCAGCTGT 2666
Qy 2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
Db 2667 TGGGCAAGGTGTGCCACGTCATCACCGAGGTGAGATCTGGAAATTTCTAGCTGAGCAGTT 2726
Qy 2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
Db 2727 CCACCTTAAACACCCATTTCCTAGTAAACAGCAGCCAAAGGACGTGCCCTACAAAGAGCT 2786
Qy 2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
Db 2787 CCTGAAGAAGCTGACCTCGCAGCAGCACGCCATTCTCATCGACCTCGGGCGAACCTTTCC 2846
Qy 2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
Db 2847 AACACATCCATACTTCTCTGCCCAGCTTGGAGCAGGTCAGCTGTCACTTTACAACATTCT 2906
Qy 2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
Db 2907 GAAGGCCTACTCGCTTCTGGACCAGGAGGTTGGATACTGCCAAGGTCTCAGCTTTGTGGC 2966
Qy 2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
Db 2967 AGGCATTTTGCTTCTTCACATGAGTGAGGAAGAGGCGTTCAAGATGCTCAAGTTCCTGAT 3026
Qy 2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
Db 3027 GTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTGCAGATCCAGAT 3086

```
Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          ||||||||| ||  || | |||| |||||||| |||||||||||| |||||||| ||
Db   3087 GTACCAGCTGTCACGGCTCCTCCACGATTACCACCGAGACCTCTACAACCACCTGGAAGA 3146

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          |||||||  ||||||     ||||| || |||||||| |||||| |||| ||||||||
Db   3147 GCACGAGACTGGCCCCCCTACGTACGCGGCTCCCTGGTTTCTCACCGTGTTCGCCTCACA 3206

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          |||||| || || || |||||||||||||||||||||| || |||||||| |||||||
Db   3207 GTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGATCAGAGGT 3266

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          |||||||||||| |||||||||| ||||| |||||||||||||||||||| ||||||||
Db   3267 CATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTACAGCATGA 3326

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
           ||||| |||||||| || ||||| ||||| | ||| || |||||||| ||| |||| ||
Db   3327 GAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGCCTGGTGCA 3386

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          |||||| |||||||||| |||||| ||||| |||||||||| || ||| | || || ||
Db   3387 GATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTCCAGGCCTA 3446

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||| || |||||||||||| ||| || || ||||| || ||||| ||||||||||||||
Db   3447 TGAGGTCGAGTACCACGTCGTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTCAGTGACAA 3506

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          |||||||||||| ||||| |||||||||||||||  || ||||||||||||||||||||
Db   3507 CCAAAGAATGGAGAAATTGGAGAAAACCAACAGCACGTTGCGCAAACAGAACCTTGACCT 3566

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
          ||| || |||||||||||||||||| |||||||||||||||||| |||||  | |||||
Db   3567 CCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACGGTAGAGAA 3626

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
           || || | ||| ||||| |||||||||||      ||| ||| | ||  |||||||
Db   3627 ACTTCTTACCAGCGAGAGTAAGCTGAAGCAGCGTGCGCTGACCCTGGAGGTGGAGCG--T 3684

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
          ||||||||||| |||||||||||||| |||  |||||   |||||||    | ||
Db   3685 CGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGCACTCCAGA 3744

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
          ||| || ||||| ||||  |||||||| ||||| |||| |||
Db   3745 GCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 3786
```

```
RESULT   4
AB029031
LOCUS       AB029031               2576 bp    mRNA    linear   PRI 04-AUG-1999
DEFINITION  Homo sapiens mRNA for KIAA1108 protein, partial cds.
ACCESSION   AB029031
VERSION     AB029031.1  GI:5689552
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hh03387.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kikuno,R., Nagase,T., Ishikawa,K., Hirosawa,M., Miyajima,N.,
            Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XIV. The complete sequences of 100 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 6 (3), 197-205 (1999)
  MEDLINE   99397452
REFERENCE   2  (bases 1 to 2576)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1. .2576
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /clone="hh03387"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1. .2292
                     /gene="KIAA1108"
     CDS             <1. .2292
                     /gene="KIAA1108"
                     /note="hh03387 cDNA clone for KIAA1108 has a 3306-bp
                     insertion."
                     /codon_start=1
                     /product="KIAA1108 protein"
                     /protein_id="BAA83060.1"
                     /db_xref="GI:5689553"
                     /translation="NSSKTKLELQKHLTTLTNQEQATIFEEVQKLRPRNEQRENELII
                     SFLRCLYEEKQKEHIHIGEMKQTSQMAAENIGSELPPSATRFRLDMLKNKAKRSLTES
                     LESILSRGNKARGLQEHSISVDLDSSLSSTLSNTSKEPSVCEKEALPISESSFKLLGS
                     SEDLSSDSESHLPEEPAPLSPQQAFRRRANTLSHFPIECQEPPQPARGSPGVSQRKLM
                     RYHSVSTETPHERKDFESKANHLGDSGGTPVKTRRHSWRQQIFLRVATPQKACDSSSR
                     YEDYSELGELPPRSPLEPVCEDGPFGPPPEEKKRTSRELRELWQKAILQQILLLRMEK
                     ENQKLQASENDLLNKRLKLDYEEITPCLKEVTTVWEKMLSTPGRSKIKFDMEKMHSAV
                     GQGVPRHHRGEIWKFLAEQFHLKHQFPSKQQPKDVPYKELLKQLTSQQHAILIDLGRT
                     FPTHPYFSAQLGAGQLSLYNILKAYSLLDQEVGYCQGLSFVAGILLLHMSEEEAFKML
                     KFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHDYHRDLYNHLEEHEIGPSLYAAPWFL
                     TMFASQFPLGFVARVFDMIFLQGTEVIFKVALSLLGSHKPLILQHENLETIVDFIKST
                     LPNLGLVQMEKTINQVFEMDIAKQLQAYEVEYHVLQEELIDSSPLSDNQRMDKLEKTN
                     SSLRKQNLDLLEQLQVANGRIQSLEATIEKLLSSESKLKQAMLTLELERSALLQTVEE
                     LRRRSAEPSDREPECTQPEPTGD"
BASE COUNT      746 a    615 c    628 g    587 t
ORIGIN
```

```
  Query Match             65.8%;  Score 2340;  DB 9;  Length 2576;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2340;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1216 aattcttccaaaacaaaactagaactgcaaaagcacctgacgacattaaccaatcaggag 1275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AATTCTTCCAAAACAAAACTAGAACTGCAAAAGCACCTGACGACATTAACCAATCAGGAG 60

Qy   1276 caggcgactatttttgaagaggttcagaaattgagaccgagaaatgagcagcgagagaat 1335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGCGACTATTTTTGAAGAGGTTCAGAAATTGAGACCGAGAAATGAGCAGCGAGAGAAT 120

Qy   1336 gaattgattatttcttttctgagatgtttatatgaagagaaacagaaagaacacatccat 1395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAATTGATTATTTCTTTTCTGAGATGTTTATATGAAGAGAAACAGAAAGAACACATCCAT 180

Qy   1396 attggggagatgaagcagacatcgcagatggcagcagagaatattggaagtgaattacca 1455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATTGGGGAGATGAAGCAGACATCGCAGATGGCAGCAGAGAATATTGGAAGTGAATTACCA 240

Qy   1456 cccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaagagatctttaaca 1515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCAGTGCCACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAGAGATCTTTAACA 300

Qy   1516 gagtctttagaaagtattttgtcccggggtaataaagccagaggcctgcaggaacactcc 1575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGTCTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTGCAGGAACACTCC 360

Qy   1576 atcagtgtggatctggatagctccctgtctagtacattaagtaacaccagcaaagagcca 1635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCAGTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCA 420

Qy   1636 tctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaagctcctcggctcc 1695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    421 TCTGTGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCC 480

Qy   1696 tcggaggacctgtccagtgactcggagagtcatctcccagaagagccagctccgctgtcg 1755
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCG 540

Qy   1756 ccccagcaggccttcaggaggcgagcaaacaccctgagtcacttccccatcgaatgccag 1815
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAG 600

Qy.  1816 gaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaacttatgaggtat 1875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTAT 660

Qy   1876 cactcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccat 1935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCAT 720

Qy   1936 cttggtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagata 1995
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATA 780

Qy   1996 ttcctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattat 2055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTAT 840

Qy   2056 tcagagctgggagagcttccccccacgatctcctttagaaccagtttgtgaagatgggccc 2115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCC 900

Qy   2116 tttggcccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaa 2175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAA 960

Qy   2176 aaggctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaa 2235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAA 1020

Qy   2236 gcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccc 2295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCC 1080

Qy   2296 tgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaa 2355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAA 1140

Qy   2356 attaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcac 2415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCATCAC 1200

Qy   2416 cgaggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcccagc 2475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGC 1260

Qy   2476 aaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcag 2535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAG 1320

Qy   2536 catgcgattcttattgaccttgggcgaacctttcctacacacccatacttctctgcccag 2595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAG 1380

Qy   2596 cttggagcaggacagctatcgctttacaacattttgaaggcctactcacttctagaccag 2655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAGACCAG 1440

Qy   2656 gaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagt 2715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTTTGCTTCTTCATATGAGT 1500

Qy   2716 gaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacag 2775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAAACAG 1560

Qy   2776 tatcggccagacatgattattttacagatccagatgtaccagctctcgaggttgcttcat 2835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCTTCAT 1620

Qy   2836 gattaccacagagacctctacaatcacctggaggagcacgagatcggccccagcctctac 2895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCTCTAC 1680

Qy   2896 gctgccccctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccaga 2955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCAGA 1740

Qy   2956 gtctttgatatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctg 3015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTCTG 1800

Qy   3016 ttgggaagccataagcccttgattctgcagcatgaaaacctagaaaccatagttgacttt 3075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACTTT 1860

Qy   3076 ataaaaagcacgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggta 3135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGGTA 1920

Qy   3136 tttgaaatggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaa 3195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TTTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTCAA 1980

Qy   3196 gaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaa 3255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGAAA 2040

Qy   3256 accaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaat 3315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCAACAGCAGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAAAT 2100

Qy   3316 ggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctg 3375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GGTAGGATCCAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGCTG 2160

Qy   3376 aagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggaggag 3435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AAGCAGGCCATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGGAG 2220

Qy   3436 ctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagccc 3495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CTGCGGCGGCGGAGCGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGCCC 2280

Qy   3496 acgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 ACGGGCGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCCTT 2340
```

```
RESULT   5
AK027355
LOCUS       AK027355               2362 bp    mRNA    linear   PRI 15-MAY-2001
DEFINITION  Homo sapiens cDNA FLJ14449 fis, clone HEMBB1001684.
ACCESSION   AK027355
VERSION     AK027355.1  GI:14041975
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens embryo, 10 weeks whole embryo, mainly body cDNA to
            mRNA, clone_lib:HEMBB1 clone:HEMBB1001684.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
            Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
```

```
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2362)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology; cDNA library construction,
            5'- & 3'-end one pass sequencing and clone selection: Helix
            Research Institute (supported by Japan Key Technology Center etc.)
            and Department of Virology, Institute of Medical Science,
            University of Tokyo
FEATURES             Location/Qualifiers
     source          1. .2362
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBB1001684"
                     /tissue_type="whole embryo, mainly body"
                     /clone_lib="HEMBB1"
                     /dev_stage="embryo, 10 weeks"
                     /note="cloning vector: pME18SFL3"
     CDS             40. .2064
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB55057.1"
                     /db_xref="GI:14041976"
                     /translation="MLKNKAKRSLTESLESILSRGNKARGLQEHSISVDLDSSLSSTL
                     SNTSKEPSVCEKEALPISESSFKLLGSSEDLSSDSESHLPEEPAPLSPQQAFRRRANT
                     LSHFPIECQEPPQPARGSPGVSQRKLMRYHSVSTETPHERKDFESKANHLGDSGGTPV
                     KTRRHSWRQQIFLRVATPQKACDSSSRYEDYSELGELPPRSPLEPVCEDGPFGPPPEE
                     KKRTSRELRELWQKAILQQILLLRMEKENQKLQASENDLLNKRLKLDYEEITPCLKEV
                     TTVWEKMLSTPGRSKIKFDMEKMHSAVGQGVPRHHRGEIWKFLAEQFHLKHQFPSKQQ
                     PKDVPYKELLKQLTSQQHAILIDLGRTFPTHPYFSAQLGAGQLSLYNILKAYSLLDQE
                     VGYCQGLSFVAGILLLHMSEEEAFKMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLL
                     HDYHRDLYNHLEEHEIGPSLYAAPWFLTMFASQFPLGFVARVFDMIFLQGTEVIFKVA
                     LSLLGSHKPLILQHENLETIVDFIKSTLPNLGLVQMEKTINQVFEMDIAKQLQAYEVE
                     YHVLQEELIDSSPLSDNQRMDKLEKTNSSLRKQNLDLLEQLQVANGRIQSLEATIEKL
                     LSSESKLKQAMLTLELERSALLQTVEELRRRSAEPSDREPECTQPEPTGD"
BASE COUNT      657 a    581 c    578 g    546 t
ORIGIN


  Query Match           59.4%;  Score 2112;  DB 9;  Length 2362;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2112;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1444 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 1503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGTGAATTACCACCCAGTGCCACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAG 60

Qy   1504 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 1563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGATCTTTAACAGAGTCTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTG 120

Qy   1564 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 1623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGGAACACTCCATCAGTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACC 180

Qy   1624 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaag 1683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCAAAGAGCCATCTGTGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAG 240

Qy   1684 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 1743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTCCTCGGCTCCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCA 300

Qy   1744 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 1803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCTCCGCTGTCGCCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCC 360.

Qy   1804 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 1863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCGAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAA 420

Qy   1864 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 1923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTTATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCC 480

Qy   1924 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 1983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAAGCAAACCATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGG 540

Qy   1984 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 2043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGGCAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGA 600

Qy   2044 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 2103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGT 660

Qy   2104 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 2163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAAGATGGGCCCTTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGA 720

Qy   2164 gagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaat 2223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGCTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAAT 780

Qy   2224 cagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaa 2283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAGAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAA 840

Qy   2284 gaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactcca 2343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GAAATTACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCA 900

Qy   2344 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 2403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGAAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTG 960

Qy   2404 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacac 2463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACAC 1020

Qy   2464 cagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctg 2523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTG 1080

Qy   2524 acttcccagcagcatgcgattcttattgaccttgggcgaaccttttcctacacacccatac 2583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACACCCATAC 1140

Qy   2584 ttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactca 2643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCA 1200

Qy   2644 cttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgctt 2703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTTTGCTT 1260

Qy   2704 cttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatgggg 2763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTTCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGG 1320

Qy   2764 ctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 2823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCG 1380

Qy   2824 aggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggc 2883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AGGTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGC 1440
```

```
Qy   2884 cccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctggga 2943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCCAGCCTCTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGA 1500

Qy   2944 ttcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtg 3003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTG 1560

Qy   3004 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 3063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCTTTAAGTCTGTTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACC 1620

Qy   3064 atagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaagacc 3123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ATAGTTGACTTTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACC 1680

Qy   3124 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtac 3183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATCAATCAGGTATTTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTAC 1740

Qy   3184 cacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggat 3243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CACGTCCTTCAAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGAT 1800

Qy   3244 aaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttg 3303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AAATTAGAGAAAACCAACAGCAGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTG 1860

Qy   3304 caggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagcagt 3363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 CAGGTGGCAAATGGTAGGATCCAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGT 1920

Qy   3364 gagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcag 3423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GAGAGCAAGCTGAAGCAGGCCATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAG 1980

Qy   3424 acggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacg 3483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACGGTGGAGGAGCTGCGGCGGCGGAGCGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACG 2040

Qy   3484 cagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatcccacac 3543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 CAGCCCGAGCCCACGGGCGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACAC 2100

Qy   3544 tgtccaggcctt 3555
          ||||||||||||
Db   2101 TGTCCAGGCCTT 2112
```

```
RESULT   6
BTA17923
LOCUS       BTA17923                1995 bp    mRNA    linear   MAM 18-NOV-1998
DEFINITION  Bos taurus mRNA for lyncein.
ACCESSION   Y17923
VERSION     Y17923.1  GI:3893076
KEYWORDS    lyn gene; lyncein.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 1995)
  AUTHORS   Xu,X.Z., Wes,P.D., Chen,H., Li,H.S., Yu,M., Morgan,S., Liu,Y. and
            Montell,C.
  TITLE     Retinal targets for calmodulin include proteins implicated in
            synaptic transmission
  JOURNAL   J. Biol. Chem. 273 (47), 31297-31307 (1998)
  MEDLINE   99030403
REFERENCE   2  (bases 1 to 1995)
  AUTHORS   Montell,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-1998) C. Montell, The Johns Hopkins University
            School of Medicine, Department of Biological Chermistry, 725 N.
            Wolfe Street, Baltimore, MD 21205-2185, USA
FEATURES             Location/Qualifiers
     source          1. .1995
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="retinal"
                     /dev_stage="adult"
     gene            2. .1672
                     /gene="lyn"
     CDS             <2. .1672
                     /gene="lyn"
                     /codon_start=1
                     /product="lyncein"
                     /protein_id="CAA76943.1"
                     /db_xref="GI:3893077"
                     /translation="ARAGVSQRKLVRYHSVSTETPHERKDFESKADHISDASRTPVKT
                     RRHSWRQQIFLRVATPQKACESPKRYEDYSELGELPPRSPLEPVCEDGPFGPVPEEKK
                     RTSHELRELWQKAILQQILLLRMEKENQKLQASENDLLNKRLKLDYEEITPCLKEVTT
                     VWEKILSTPGRSKIKFDMEKMHSAVGQGVPRHHRGEIWKFLAEQYHLKHPFPCKQQPK
                     DTPYKELLKQLTSQQHAILIDLGRTFPTHPYYSAQLGAGQLSLYNILKAYSLLDQEVG
                     YCQGLSFVAGILLLHMGEEEAFNMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHD
                     YHRDLYNHLEEHEIGPSLYAAPWFLTVFASQFPLGFVARVFDMIFLQGSEVIFKVALS
                     LLGSHKPLILQHENLETIVDFIKSTLPNLGLVQMEKTISQVFETDISKQLQAYEVEYH
                     VLQEELIDSSPLSDNQRMDKLEKTNSSLRKQNLDLLEQLQVANGRIQSLEATVEKLLT
                     SESKLKQATLALELERSALLQTVEQLRRQTAELGSQESDPTLPKPSGD"
BASE COUNT      560 a    516 c    497 g    422 t
ORIGIN
```

```
  Query Match             38.2%;  Score 1359.2;  DB 4;  Length 1995;
  Best Local Similarity   88.4%;  Pred. No. 1.9e-286;
  Matches 1475;  Conservative    0;  Mismatches  193;  Indels    0;  Gaps    0;

Qy   1844 cgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatg 1903
          | |||||  ||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      9 CAGGGGTCTCGCAAAGGAAACTTGTGAGGTATCACTCAGTGAGCACAGAGACGCCTCATG 68

Qy   1904 aacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtga 1963
          ||||||||||||||| ||||||||  ||||  |  | ||| |  |  |||| |||||||
Db     69 AACGAAAGGACTTTGAGTCCAAAGCCGACCACATCAGCGATGCCAGCAGGACCCCTGTGA 128

Qy   1964 agacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaagg 2023
          |||| ||||| || ||  |||||||||||||| ||||| || || |||||||| |||||||
Db    129 AGACACGGAGACACTCGTGGAGGCAGCAGATCTTCCTTCGGGTGGCCACCCCCCAGAAGG 188

Qy   2024 cgtgcgattcttccagcagatatgaagattattcagagctgggagagcttcccccacgat 2083
          | ||||| ||  |||  ||||| |||||||| |||||| ||||||||||||||||||||
Db    189 CCTGCGAGTCCCCCAAAAGATACGAAGATTACTCAGAGTTGGGAGAGCTTCCCCCACGAT 248

Qy   2084 ctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaaa 2143
          ||||  |||||||||||||||||||| ||||||||||    ||||||||||||||||||
Db    249 CTCCACTAGAACCAGTTTGTGAAGATGGACCCTTTGGCCCGGTCCCAGAGGAAAAGAAAA 308

Qy   2144 ggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgc 2203
          | |||||| ||| || || |||||||||||||||| ||||| ||||||||| ||||||
Db    309 GAACATCTCATGAACTTCGGGAGCTGTGGCAAAAGGCCATTCTACAACAGATATTGCTGC 368

Qy   2204 ttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagc 2263
          | || |||||||| ||||| || || || |||||||||| ||||||||||||||||||||
Db    369 TCAGGATGGAGAAAGAAAACCAAAAACTGCAAGCCTCTGAGAATGATTTGCTGAACAAGC 428

Qy   2264 gcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtggg 2323
          |||| ||||| |||||||||||||| ||||||||||||||||||||||||||||||||||
Db    429 GCCTAAAGCTTGATTATGAAGAAATCACTCCCTGTCTTAAAGAAGTGACTACAGTGTGGG 488

Qy   2324 aaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcact 2383
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    489 AAAAGATACTTAGCACTCCAGGAAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACT 548

Qy   2384 cggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctg 2443
          |||||||||| |||||||||| || || ||||| |||||||||||||||||  | || |
Db    549 CGGCTGTTGGCCAAGGTGTGCCGCGACACCACCGGGGTGAAATCTGGAAATTCTTGGCGG 608
```

```
Qy   2444 agcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccataca 2503
          |||| | |||||| || |||| |||||||| |||||||||||| |||||   ||| ||||
Db    609 AGCAGTACCACCTGAAGCACCCGTTTCCCTGCAAACAGCAGCCCAAGGACACGCCCTACA 668

Qy   2504 aagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaa 2563
          |||| |||||||||||||| || |||||||||| || ||||| || ||||||||||||
Db    669 AAGAGCTCTTAAAGCAGCTCACCTCCCAGCAGCACGCCATTCTCATCGACCTTGGGCGAA 728

Qy   2564 cctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttaca 2623
          ||||||| ||||| || |||| |||||| ||||| |||||||| ||||| || ||||||||
Db    729 CCTTTCCAACACATCCGTACTACTCTGCTCAGCTAGGAGCAGGGCAGCTGTCACTTTACA 788

Qy   2624 acattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagct 2683
          ||||  |||||||||||||| |||||||||| |||||||| ||||||||||||||||||
Db    789 ACATCCTGAAGGCCTACTCGCTTCTAGACCAAGAAGTGGGCTATTGCCAAGGTCTCAGCT 848

Qy   2684 ttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagt 2743
          ||||||||||||   |||| |||||||| |||| ||||||||||||| ||||||||||
Db    849 TTGTAGCAGGCATCCTGCTGCTTCATATGGGTGAAGAAGAGGCGTTTAACATGCTCAAGT 908

Qy   2744 ttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacaga 2803
          |||||||||||||||||| |||||||||| |||||||| ||||||||||||||||||||
Db    909 TTCTGATGTTTGACATGGGACTGCGGAAACAATATCGGCCGGACATGATTATTTTACAGA 968

Qy   2804 tccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacc 2863
          |||||||||||||||||||||| |||| || || |||||||| |||||||||| ||||
Db    969 TCCAGATGTACCAGCTCTCGAGGCTGCTCCACGACTACCACAGGGACCTCTACAACCACC 1028

Qy   2864 tggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttg 2923
          |||||||||| ||||| |||||||||||||| ||||| |||||||||||||| ||||||
Db   1029 TGGAGGAGCATGAGATTGGCCCCAGCCTCTATGCTGCGCCCTGGTTCCTCACCGTGTTTG 1088

Qy   2924 cctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaa 2983
          |||||||||||| ||||| |||| |||||||||||||||||| ||||||||||||
Db   1089 CCTCACAGTTCCCCCTGGGCTTCGTCGCCAGAGTCTTTGATATGATCTTTCTTCAGGGAT 1148

Qy   2984 cagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctgc 3043
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1149 CAGAGGTCATATTTAAAGTGGCTTTAAGTCTGCTGGGAAGCCATAAGCCCTTGATTCTGC 1208

Qy   3044 agcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggct 3103
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||
Db   1209 AGCATGAAAACCTAGAAACCATAGTTGACTTTATAAAAAGCACACTACCCAACCTGGGCT 1268

Qy   3104 tggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttac 3163
          |||| |||||||| ||||| |||| |||||| ||||| | ||||||| | ||||||||||
Db   1269 TGGTGCAGATGGAGAAGACTATCAGTCAGGTGTTTGAGACGGACATCTCGAAACAGTTAC 1328

Qy   3164 aagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctca 3223
          |||||||||||||||||||||||| ||||| ||||| |||||||||| ||||||||||
Db   1329 AAGCTTATGAAGTTGAGTACCACGTGCTTCAGGAAGAGCTTATCGATTCTTCTCCTCTCA 1388

Qy   3224 gtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaacc 3283
          |||||||||||||||||| ||| ||||||| |||||||||||||||||||||||||||
Db   1389 GTGACAACCAAAGAATGGACAAACTAGAGAAGACCAACAGCAGCTTACGCAAACAGAACC 1448

Qy   3284 ttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccacca 3343
          |||||||||||||||| |||||| || ||||| |||||||| ||||| || |||||
Db   1449 TTGACCTCCTTGAACAGCTGCAGGTCGCCAATGGGAGGATCCAGAGCCTGGAAGCCACTG 1508

Qy   3344 ttgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactgg 3403
          ||||||||||||||  || |||||||||||||||||| |  ||  |||| || ||||
Db   1509 TTGAGAAGCTCCTGACTAGCGAGAGCAAGCTGAAGCAGGCGACCCTGGCCTTGGAGCTGG 1568

Qy   3404 agcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcg 3463
          || |||| |||||||||||||||||||| |||| ||| ||| ||  |||||||   ||
Db   1569 AGAGGTCAGCCCTGCTGCAGACGGTGGAGCAGCTCCGGAGGCAGACGGCAGAGCTGGGCA 1628

Qy   3464 accgggagcctgagtgcacgcagcccgagcccacgggcgactgacagc 3511
            || |||| ||||   ||| | |||| ||||| | || ||||||  ||
Db   1629 GCCAGGAGTCTGACCCCACACTGCCCAAGCCCTCAGGTGACTGATGGC 1676
```

```
RESULT   7
BC004675
LOCUS       BC004675               3051 bp    mRNA    linear   ROD 19-JUL-2001
DEFINITION  Mus musculus, clone IMAGE:3500261, mRNA, partial cds.
ACCESSION   BC004675
VERSION     BC004675.1  GI:14919386
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3051)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gibbs,R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 9 Row: m Column: 4.
FEATURES             Location/Qualifiers
     source          1. .3051
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3500261"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             <1. .1485
                     /codon_start=1
                     /product="Unknown (protein for IMAGE:3500261)"
                     /protein_id="AAH04675.1"
                     /db_xref="GI:14919387"
                     /translation="ACDSPSRYEDYSELGELPPRSPLEPVCEDGPFGPVQEEKRKTSR
                     ELRELWKKAILQQILLLRMEKENQKLQASENDLLNKRLKLDYDEITPCLKEVTTVWEK
                     MLSTPGRSKIKFDMEKVHSAVGQGVPRHHRGEIWKFLAEQFHLKHPFPSKQQPKDVPY
                     KELLKKLTSQQHAILIDLGRTFPTHPYFSAQLGAGQLSLYNILKAYSLLDQEVGYCQG
                     LSFVAGILLLHMSEEEAFKMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHDYHRD
                     LYNHLEEHEIGPSLYAAPWFLTVFASQFPLGFVARVFDMIFLQGSEVIFKVALSLLGS
                     HKPLILQHENLETIVDFIKNTLPNLGLVQMEKTISQVFEMDIAKQLQPYEVEYHVLQE
                     ELIESSPLSDNQRMEKLEKTNSSLRKQNLDLLEQLQVANARIQSLEATVEKLLTSESK
                     LKQAALTLEVERSALLQMVEELRRQSARPSTPEPDCTQLEPTGD"
BASE COUNT      817 a    715 c    756 g    763 t
ORIGIN
```

```
  Query Match             33.2%;  Score 1179.6;  DB 10;  Length 3051;
  Best Local Similarity   87.0%;  Pred. No. 3.5e-247;
  Matches 1296;  Conservative    0;  Mismatches  194;  Indels    0;  Gaps    0;

Qy   2023 gcgtgcgattcttccagcagatatgaagattattcagagctgggagagctttcccccacga 2082
          || || || ||  | |||||||||||||||||| |||||||||||||| || |||||
Db      1 GCTTGTGACTCCCCGAGCAGATATGAAGATTATTCCGAGCTGGGAGAGCTCCCTCCACGC 60

Qy   2083 tctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaa 2142
```

```
Db      61 TCCCCTTTAGAACCGGTGTGTGAGGACGGCCCATTTGGCCCAGTACAGGAAGAAAAGAGG 120

Qy    2143 aggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctg 2202

Db     121 AAGACGTCACGCGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTGCAGCAGATCCTGCTG 180

Qy    2203 cttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaag 2262

Db     181 CTCAGGATGGAGAAGGAGAATCAGAAGCTACAAGCCTCTGAAAACGATTTGCTGAACAAA 240

Qy    2263 cgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgg 2322

Db     241 CGCCTCAAGCTTGACTATGATGAAATCACTCCGTGTCTTAAAGAAGTCACTACAGTGTGG 300

Qy    2323 gaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcac 2382

Db     301 GAAAAGATGCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAAGTGCAC 360

Qy    2383 tcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagct 2442

Db     361 TCAGCTGTTGGGCAAGGTGTGCCACGTCATCACCGAGGTGAGATCTGGAAATTTCTAGCT 420

Qy    2443 gagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatac 2502

Db     421 GAGCAGTTCCACCTTAAACACCCATTTCCTAGTAAACAGCAGCCAAAGGACGTGCCCTAC 480

Qy    2503 aaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcga 2562

Db     481 AAAGAGCTCCTGAAGAAGCTGACCTCGCAGCAGCACGCCATTCTCATCGACCTCGGGCGA 540

Qy    2563 acctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttac 2622

Db     541 ACCTTTCCAACACATCCATACTTCTCTGCCCAGCTTGGAGCAGGTCAGCTGTCACTTTAC 600

Qy    2623 aacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagc 2682

Db     601 AACATTCTGAAGGCCTACTCGCTTCTGGACCAGGAGGTTGGATACTGCCAAGGTCTCAGC 660

Qy    2683 tttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaag 2742

Db     661 TTTGTGGCAGGCATTTTGCTTCTTCACATGAGTGAGGAAGAGGCGTTCAAGATGCTCAAG 720

Qy    2743 tttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacag 2802

Db     721 TTCCTGATGTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTGCAG 780

Qy    2803 atccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcac 2862

Db     781 ATCCAGATGTACCAGCTGTCACGGCTCCTCCACGATTACCACCGAGACCTCTACAACCAC 840

Qy    2863 ctggaggagcacgagatcggcccccagcctctacgctgccccctggttcctcaccatgttt 2922

Db     841 CTGGAAGAGCACGAGATCGGCCCCAGCCTCTACGCGGCTCCCTGGTTTCTCACCGTGTTC 900

Qy    2923 gcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcaggga 2982

Db     901 GCCTCACAGTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGA 960

Qy    2983 acagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctg 3042

Db     961 TCAGAGGTCATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTA 1020

Qy    3043 cagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggc 3102

Db    1021 CAGCATGAGAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGC 1080

Qy    3103 ttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagtta 3162

Db    1081 CTGGTGCAGATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTC 1140

Qy    3163 caagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctc 3222

Db    1141 CAGCCCTATGAGGTCGAGTACCACGTGCTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTC 1200

Qy    3223 agtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaac 3282

Db    1201 AGTGACAACCAAAGAATGGAGAAATTGGAGAAAACCAACAGCAGCTTGCGCAAACAGAAC 1260

Qy    3283 cttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccacc 3342

Db    1261 CTTGACCTCCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACG 1320

Qy    3343 attgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactg 3402

Db    1321 GTAGAGAAACTTCTTACCAGCGAGAGTAAGCTGAAGCAGGCTGCGCTGACCCTGGAGGTG 1380

Qy    3403 gagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagc 3462

Db    1381 GAGCGCTCCGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGC 1440

Qy    3463 gaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512

Db    1441 ACTCCAGAGCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 1490
```

RESULT 8
AK057182
LOCUS AK057182 1788 bp mRNA linear PRI 31-OCT-2001
DEFINITION Homo sapiens cDNA FLJ32620 fis, clone STOMA2000386, highly similar to Bos taurus mRNA for lyncein.
ACCESSION AK057182
VERSION AK057182.1 GI:16552778
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens stomach cDNA to mRNA, clone_lib:STOMA2 clone:STOMA2000386.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1788)
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES Location/Qualifiers
source 1. .1788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STOMA2000386"
/tissue_type="stomach"
/clone_lib="STOMA2"
/note="cloning vector: pME18SFL3"

```
BASE COUNT      481 a    412 c    434 g    461 t
ORIGIN

  Query Match           28.1%;  Score 998.4;  DB 9;  Length 1788;
  Best Local Similarity  99.9%;  Pred. No. 1.3e-207;
  Matches  999;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2556 tgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatc 2615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    553 TGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATC 612

Qy   2616 gctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaagg 2675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    613 GCTTTACAACATTTTGAAGGCCTACTCACTTCTAGACCAGGAAGTGGGATATTGCCAAGG 672

Qy   2676 tctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaat 2735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    673 TCTCAGCTTTGTAGCAGGCATTTTGCTTCTTCATATGAGTGAGGAAGAGGCGTTTAAAAT 732

Qy   2736 gctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattat 2795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    733 GCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTAT 792

Qy   2796 tttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctcta 2855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    793 TTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCTTCATGATTACCACAGAGACCTCTA 852

Qy   2856 caatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcac 2915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    853 CAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCTCTACGCTGCCCCCTGGTTCCTCAC 912

Qy   2916 catgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttct 2975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    913 CATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCAGAGTCTTTGATATGATTTTTCT 972

Qy   2976 tcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccctt 3035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    973 TCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTCTGTTGGGAAGCCATAAGCCCTT 1032

Qy   3036 gattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaa 3095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1033 GATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACTTTATAAAAAGCACGCTACCCAA 1092

Qy   3096 ccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaa 3155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1093 CCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGGTATTTGAAATGGACATCGCTAA 1152

Qy   3156 acagttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctc 3215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1153 ACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTCAAGAAGAACTTATCGATTCCTC 1212

Qy   3216 tcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaa 3275
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1213 TCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGAAGACCAACAGCAGCTTACGCAA 1272

Qy   3276 acagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttga 3335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1273 ACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAAATGGTAGGATCCAAAGCCTTGA 1332

Qy   3336 ggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttacctt 3395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1333 GGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGCTGAAGCAGGCCATGCTTACCTT 1392

Qy   3396 agaactggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcaga 3455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1393 AGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGGAGCTGCGGCGGCGGAGCGCAGA 1452

Qy   3456 gcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctg 3515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1453 GCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGCCCACGGGCGACTGACAGCTCTG 1512

Qy   3516 caggagagattgcaacaccatcccacactgtccaggcctt 3555
          |||||||||||||||||||||||||||||||||||||||
Db   1513 CAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCCTT 1552

RESULT   9
AB011175
LOCUS       AB011175                5922 bp    mRNA    linear   PRI 10-APR-1998
DEFINITION  Homo sapiens mRNA for KIAA0603 protein, complete cds.
ACCESSION   AB011175
VERSION     AB011175.1  GI:3043729
KEYWORDS    KIAA0603 protein.
SOURCE      Homo sapiens male brain cDNA to mRNA, clone_lib:pBluescriptII SK
            plus clone:HG1488b.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5922)
  AUTHORS   Ohara,O., Nagase,T. and Ishikawa,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
            DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
REFERENCE   2  (sites)
  AUTHORS   Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
            Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes. IX.
            The complete sequences of 100 new cDNA clones from brain which can
            code for large proteins in vitro
  JOURNAL   DNA Res. 5 (1), 31-39 (1998)
  MEDLINE   98290545
FEATURES             Location/Qualifiers
     source          1. .5922
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HG1488b"
                     /sex="male"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            348. .4247
                     /gene="KIAA0603"
     CDS             348. .4247
                     /gene="KIAA0603"
                     /codon_start=1
                     /product="KIAA0603 protein"
                     /protein_id="BAA25529.1"
                     /db_xref="GI:3043730"
                     /translation="MEPPSCIQDEPFPHPLEPEPGVSAQPGPGKPSDKRFRLWYVGGS
                     CLDHRTTLPMLPWLMAEIRRRSQKPEAGGCGAPAARE VILVLSAPFLRCVPAPGAGAS
                     GGTSPSATQPNPAVFIFEHKAQHISRFIHNSHDLTYFAYLIKAQPDDPESQMACHVFR
                     ATDPSQVPDVISSIRQLSKAAMKEDAKPSKDNEDAFYNSQKFEVLYCGKVTVTHKKAP
                     SSLIDDCMEKFSLHEQQRLKIQGEQRGPDPGEDLADLEVVVPGSPGDCLPEEADGTDT
                     HLGLPAGASQPALTSSRVCFPERILEDSGFDEQQEFRSRCSSVTGVQRRVHEGSQKSQ
                     PRRRHASAPSHVQPSDSEKNRTMLFQVGRFEINLISPDTKSVVLEKNFKDISSCSQGI
                     KHVDHFGFICRESPEPGLSQYICYVFQCASESLVDEVMLTLKQAFSTAAALQSAKTQI
                     KLCEACPMHSLHKLCERIEGLYPPRAKLVIQRHLSSLTDNEQADIFERVQKMKPVSDQ
                     EENELVILHLRQLCEAKQKTHVHIGEGPSTISNSTIPENATSSGRFKLDILKNKAKRS
                     LTSSLENIFSRGANRMRGRLGSVDSFERSNSLASEKDYSPGDSPPGTPPASPPSSAWQ
                     TFPEEDSDSPQFRRRAHTFSHPPSSTKRKLNLQDGRAQGVRSPLLRQSSSEQCSNLSS
                     VRRMYKESNSSSSLPSLHTSFSAPSFTAPSFLKSFYQNSGRLSPQYENEIRQDTASES
                     SDGEGRKRTSSTCSNESLSVGGTSVTPRRISWRQRIFLRVASPMNKSPSAMQQQDGLD
                     RNELLPLSPLSPTMEEEPLVIFLSGEDDPEKIEERKKSKELRSLWRKAIHQQILLLRM
                     EKENQKLEGASRDELQSRKVKLDYEEVGACQKEVLITWDKKLLNCRAKIRCDMEDIHT
                     LLKEGVPKSRRGEIWQFLALQYRLRHRLPNKQQPPDISYKELLKQLTAQQHAILVDLG
                     RTFPTHPYFSVQLGPGQLSLFNLLKAYSLLDKEVGYCQGISFVAGVLLLHMSEEQAFE
                     MLKFLMYDLGFRKQYRPDMMSLQIQMYQLSRLLHDYHRDLYNHLEENEISPSLYAAPW
                     FLTLFASQFSLGFVARVFDIIFLQGTEVIFKVALSLLSSQETLIMECESFENIVEFLK
                     NTLPDMNTSEMEKIITQVFEMDISKQLHAYEVEYHVLQDELQESSYSCEDSETLEKLE
                     RANSQLKRQNMDLLEKLQVAHTKIQALESNLENLLTRETKMKSLIRTLEQEKMAYQKT
                     VEQLRKLLPADALANCDLLLRDLNCNPNNKAKIGNKP"
BASE COUNT     1624 a   1408 c   1432 g   1458 t
ORIGIN
```

Query Match 19.2%; Score 682.2; DB 9; Length 5922;
Best Local Similarity 56.6%; Pred. No. 1.9e-138;
Matches 1585; Conservative 0; Mismatches 1028; Indels 186; Gaps 10;

```
Qy    823 agcggacacaatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacg 882
Db   1368 AGCGCACCCAGTCACGTCCAGCCCTCGGACTCGGAGAAGAACAGGACCATGCTCTTCCAG 1427

Qy    883 attggccagtctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggagaaa 942
Db   1428 GTTGGGCGATTTGAGATTAACCTTATCAGTCCAGACACTAAATCAGTTGTGCTAGAAAAG 1487

Qy    943 aattttaaggagatatccttttgctctcagggcatcagacacgtggaccactttgggttt 1002
Db   1488 AATTTTAAAGATATCTCCTCTTGTTCTCAGGGTATAAAGCATGTCGATCACTTTGGCTTT 1547

Qy   1003 atctgtcgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgc 1062
Db   1548 ATCTGCCGGGAGTCTCCAGAGCCTGGACTTAGCCAGTATATTTGTTATGTATTCCAGTGT 1607

Qy   1063 acaaatgaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggcc 1122
Db   1608 GCCAGCGAATCTCTGGTTGATGAGGTAATGCTGACTCTGAAACAGGCCTTCAGTACGGCG 1667

Qy   1123 gcagtgcagcagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagc 1182
Db   1668 GCTGCCCTGCAGAGTGCCAAGACGCAGATTAAACTGTGTGAGGCCTGCCCGATGCACTCT 1727

Qy   1183 ctgcacaagctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactg 1242
Db   1728 TTGCATAAGCTCTGTGAAAGGATTGAAGGTCTCTACCCACCAAGAGCCAAGCTGGTGATA 1787

Qy   1243 caaaagcacctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcag 1302
Db   1788 CAGAGGCATCTCTCATCACTGACAGATAATGAGCAAGCTGACATCTTTGAAAGAGTTCAG 1847

Qy   1303 aaattgagaccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgt 1362
Db   1848 AAAATGAAGCCAGTCAGTGACCAGGAAGAAAATGAACTTGTGATTTTACACCTGAGGCAG 1907

Qy   1363 ttatatgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcgcag 1422
Db   1908 CTGTGTGAAGCCAAGCAGAAGACACACGTGCACATCGGGG---------AAGGCCCTTCT 1958

Qy   1423 atggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggctagat 1482
Db   1959 ACTATTTCAAATAGTACAATCCCAGAAAATGCAACAAGCAGTGGAAGGTTCAAACTTGAC 2018

Qy   1483 atgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccgg 1542
Db   2019 ATTCTGAAAAATAAAGCTAAGAGATCCTTAACTAGCTCCCTGGAAAATATCTTCTCAAGG 2078

Qy   1543 gg---taataaagccagaggcctgcaggaacactccatcagt----------------- 1581
Db   2079 GGAGCTAACAGAATGAGAGGTCGGCTTGGAAGTGTGGACAGTTTTGAACGGTCCAACAGT 2138

Qy   1582 ------------------------------------------------------------ 1581
Db   2139 CTTGCTTCAGAGAAGGACTACTCACCAGGGGATTCTCCACCAGGGACACCGCCAGCGTCC 2198

Qy   1582 ----------------------------gtggatctggatagctccctgtctagt 1608
Db   2199 CCACCGTCCTCAGCTTGGCAAACGTTTCCCGAAGAGGATTCCGACTCCCCGCAGTTTCGA 2258

Qy   1609 acattaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctct 1668
Db   2259 AGACGGGCACACACGTTCAGCCACCCACCTTCAAGCACAAAGAGAAAGCTGAATTTGCAG 2318

Qy   1669 gagagctcctttaagctcctcggctcctcg----------------------------gagg 1702
Db   2319 GATGGGAGGGCTCAGGGTGTGCGTTCCCCTCTGCTGAGGCAGAGCTCCAGTGAACAGTGC 2378

Qy   1703 acctgtccagtgactcggagagtcatctcccagaagagccagctccgctgtcgccccagc 1762
Db   2379 AGCAATCTTTCGTCAGTTCGACGCATGTACAAGGAGAGTAATTCTTCCTCCAGTCTTCCA 2438

Qy   1763 aggccttcaggaggcgagcaaacaccctgagtcacttccccatcgaatgccaggaacctc 1822
Db   2439 AGTCTTCACACTTCCTTCTCTGCCCCTTCCTTCACTGCCCCCTCTTTCCTGAAAAGCTTT 2498

Qy   1823 cacaacctgcccgggggtc---cccgggggtttcgcaaaggaaacttatgaggtatcact 1879
Db   2499 TACCAGAATTCAGGTAGACTGTCCCCACAGTATGAAAATGAAATCAGACAAGACACTGCT 2558

Qy   1880 cagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatcttg 1939
Db   2559 TCAGAATCAAGTGATGGAGAAGGGAGAAAAAGGACCTCATCTACCTGCAGCAATGAGTCC 2618

Qy   1940 gtgattct-ggtgggactcctgtgaagacccggaggcattcctggaggcagcagatattc 1998
Db   2619 CTAAGTGTGGGAGGAACCTCTGTCACTCCTCGCCGGATCTCCTGGCGGCAGCGCATTTTC 2678

Qy   1999 ctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattca 2058
Db   2679 CTCAGGGTTGCTTCTCCCATGAACAAATCTCCCTCAGCAATGCAACAGCAAGATGGATTG 2738

Qy   2059 gagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgggcccttt- 2117
Db   2739 GACAGGAACGAGCTGCTGCCACTGTCCCCCCTCTCTCCAACCATGGAGGAGGAACCGCTG 2798

Qy   2118 ----------------------tggcccccccaccagaggaaaagaaaaggacatctcgt 2154
Db   2799 GTTATATTCCTGTCTGGGGAGGATGACCCAGAAAAGATTGAAGAAAGAAAGAAATCAAAA 2858

Qy   2155 gagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggag 2214
Db   2859 GAACTGAGGAGCTTGTGGAGAAAAGCTATACACCAACAAATCTTGTTACTTCGAATGGAA 2918

Qy   2215 aaggaaaatcagaagctcca---agcctctgaaaatgatttgctgaacaagcgcctgaag 2271
Db   2919 AAAGAAAACCAGAAACTTGAAGGAGCAAGCAGAGATGAACTCCAGTCCAGAAAAGTTAAA 2978

Qy   2272 ctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatg 2331
Db   2979 TTAGACTATGAAGAAGTTGGTGCATGTCAGAAAGAGGTCTTAATAACTTGGGATAAGAAG 3038

Qy   2332 cttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgtt 2391
Db   3039 TTGTTAA---ACTGCAGAGCTAAAATCAGATGTGATATGGAAGATATTCATACTCTTCTT 3095

Qy   2392 gggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattc 2451
Db   3096 AAAGAAGGAGTTCCCAAAAGTCGACGAGGAGAAATTTGGCAGTTTCTGGCTTTACAGTAC 3155

Qy   2452 caccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactc 2511
Db   3156 CGACTCAGACACAGATTGCCTAATAAACAACAGCCTCCTGACATATCCTATAAGGAACTT 3215

Qy   2512 ttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcct 2571
Db   3216 TTGAAGCAGCTCACTGCTCAGCAGCATGCGATTCTCGTGGATTTAGGAAGGACGTTTCCT 3275

Qy   2572 acacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacattttg 2631
Db   3276 ACTCACCCTTACTTTTCAGTACAGCTTGGGCCAGGACAGCTGTCACTGTTTAACCTCCTG 3335

Qy   2632 aaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagca 2691
Db   3336 AAAGCCTATTCTTTGCTGGACAAAGAAGTGGGATACTGTCAGGGGATCAGCTTTGTGGCT 3395

Qy   2692 ggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatg 2751
Db   3396 GGAGTCCTGCTTCTGCACATGAGTGAAGAGCAAGCCTTTGAAATGCTGAAATTCCTCATG 3455

Qy   2752 tttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagatg 2811
```

```
Db    3456 TATGACCTCGGCTTCCGCAAGCAGTACAGACCTGACATGATGTCGCTGCAGATTCAAATG 3515

Qy    2812 taccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggaggag 2871

Db    3516 TACCAGCTGTCCAGGCTCCTTCATGACTATCACAGAGATCTCTACAATCACCTTGAAGAA 3575

Qy    2872 cacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacag 2931

Db    3576 AATGAAATCAGCCCCAGTCTTTATGCTGCCCCCTGGTTCCTCACATTGTTTGCCTCTCAG 3635

Qy    2932 ttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtc 2991

Db    3636 TTTTCATTAGGATTTGTAGCCAGAGTTTTTGATATTATTTTTCTTCAGGGAACTGAAGTT 3695

Qy    2992 atatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaa 3051

Db    3696 ATATTCAAGGTTGCACTCAGCCTACTGAGCAGCCAAGAGACACTTATAATGGAATGTGAG 3755

Qy    3052 aacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtacag 3111

Db    3756 AGCTTTGAAAATATTGTTGAGTTTCTTAAAAACACGCTACCTGATATGAATACCTCTGAA 3815

Qy    3112 atggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttat 3171

Db    3816 ATGGAAAAAATTATTACCCAGGTTTTTGAGATGGATATTTCTAAGCAGTTGCATGCCTAT 3875

Qy    3172 gaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagt---gac 3228

Db    3876 GAGGTGGAATATCATGTGCTACAGGATGAGCTTCAGGAATCTTCATATTCCTGTGAGGAT 3935

Qy    3229 aaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgac 3288

Db    3936 AGTGAAACTTTGGAGAAGCTGGAGAGGGCCAATAGCCAACTGAAAAGACAAAACATGGAC 3995

Qy    3289 ctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgag 3348

Db    3996 CTCCTAGAAAAATTACAGGTAGCTCATACTAAAATCCAGGCCTTGGAATCAAACCTGGAA 4055

Qy    3349 aagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcgg 3408

Db    4056 AATCTTTTGACGAGAGAGACCAAAATGAAGTCTTTAATCCGGACCCTGGAACAAGAAAAA 4115

Qy    3409 tcggccctgctgcagacggtggaggagctgcggcggcgg 3447

Db    4116 ATGGCTTATCAAAAGACAGTGGAGCAACTCCGGAAGCTG 4154


RESULT  10
AC009595/c
ID   AC009595   standard; DNA; HTG; 189811 BP.
XX
AC   AC009595;
XX
SV   AC009595.3
XX
DT   30-AUG-1999 (Rel. 60, Created)
DT   03-JUL-2000 (Rel. 64, Last updated, Version 4)
XX
DE   Homo sapiens chromosome 4 clone RP11-392K14 map 4, WORKING DRAFT SEQUENCE,
DE   28 unordered pieces.
XX
KW   HTG; HTGS_DRAFT; HTGS_PHASE1.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-189811
RA   Birren B., Linton L., Nusbaum C., Lander E.;
RT   "Homo sapiens chromosome 4, clone RP11-392K14";
RL   Unpublished.
XX
RN   [2]
RP   1-189811
RA   Birren B., Linton L., Nusbaum C., Lander E., Allen N., Anderson M.,
RA   Baker J., Baldwin J., Barna N., Beckerly R., Benn J., Brown A., Castle A.,
RA   Cerny J., Colangelo M., Collins S., Collymore A., Cooke P., DeArellano K.,
RA   Depayre E., Devon K., Dewar K., Donelan L., Doyle M., Ferreira P.,
RA   FitzHugh W., Forrest C., Funke R., Gage D., Galagan J., Gardyna S.,
RA   Gilbert D., Grant G., Hagos B., Heaford A., Horton L., Howland J.C.,
RA   Jones C., Kann L., Karatas A., Lehoczky J., Lieu C., Locke K.,
RA   Macdonald P., Marquis N., McEwan P., McGurk A., McKernan K., McLaughlin J.,
RA   Meldrim J., Molla M., Morris W., Morrow J., Mychaleckyj J., Naylor J.,
RA   Niloff M., O'Connor T., O'Donnell P., Pavlin B., Peterson K., Pollara V.,
RA   Riley R., Roberts D., Roy A., Severy P., Stange-Thomann N., Stojanovic N.,
RA   Stone C., Subramanian A., Tesfaye S., Torruella-Miller I., Vassiliev H.,
RA   Vo A., Wagner A., Wheeler J., Wu X., Wyman D., Ye W.J., Zody M.;
RT   ;
RL   Submitted (27-AUG-1999) to the EMBL/GenBank/DDBJ databases.
RL   Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,
RL   Cambridge, MA 02141, USA
XX
CC   On Jun 23, 2000 this sequence version replaced gi:6006228.
CC   All repeats were identified using RepeatMasker:
CC   Smit, A.F.A. & Green, P. (1996-1997)
CC   http://ftp.genome.washington.edu/RM/RepeatMasker.html
CC   -------------- Genome Center
CC   Center: Whitehead Institute/ MIT Center for Genome Research
CC   Center code: WIBR
CC   Web site: http://www-seq.wi.mit.edu
CC   Contact: sequence_submissions@genome.wi.mit.edu
CC   -------------- Project Information
CC   Center project name: L1755
CC   Center clone name: 392_K_14
CC   -------------- Summary Statistics
CC   Sequencing vector: M13; M77815; 100% of reads
CC   Chemistry: Dye-primer-amersham; 5% of reads
CC   Chemistry: Dye-terminator Big Dye; 95% of reads
CC   Assembly program: Phrap; version 0.960731
CC   Consensus quality: 164704 bases at least Q40
CC   Consensus quality: 175594 bases at least Q30
CC   Consensus quality: 181024 bases at least Q20
CC   Insert size: 214000; agarose-fp
CC   Insert size: 187111; sum-of-contigs
CC   Quality coverage: 3.0 in Q20 bases; agarose-fp
CC   Quality coverage: 3.5 in Q20 bas.
CC   * NOTE: This is a 'working draft' sequence. It currently
CC   * consists of 28 contigs. The true order of the pieces
CC   * is not known and their order in this sequence record is
CC   * arbitrary. Gaps between the contigs are represented as
CC   * runs of N, but the exact sizes of the gaps are unknown.
CC   * This record will be updated with the finished sequence
CC   * as soon as it is available and the accession number will
CC   * be preserved.
CC   *        1      451: contig of 451 bp in length
CC   *      452 551: gap of      100 bp
CC   *      552     2663: contig of 2112 bp in length
CC   *     2664 2763: gap of      100 bp
CC   *     2764     4124: contig of 1361 bp in length
CC   *     4125 4224: gap of      100 bp
CC   *     4225     6767: contig of 2543 bp in length
CC   *     6768 6867: gap of      100 bp
CC   *     6868     8750: contig of 1883 bp in length
CC   *     8751 8850: gap of      100 bp
CC   *     8851    11272: contig of 2422 bp in length
CC   *    11273 11372: gap of      100 bp
CC   *    11373    13411: contig of 2039 bp in length
CC   *    13412 13511: gap of      100 bp
CC   *    13512    16962: contig of 3451 bp in length
CC   *    16963 17062: gap of      100 bp
CC   *    17063    22544: contig of 5482 bp in length
CC   *    22545 22644: gap of      100 bp
CC   *    22645    26284: contig of 3640 bp in length
CC   *    26285 26384: gap of      100 bp
```

```
CC   *    26385    31223: contig of 4839 bp in length
CC   *    31224 31323: gap of     100 bp
CC   *    31324    36014: contig of 4691 bp in length
CC   *    36015 36114: gap of     100 bp
CC   *    36115    41939: contig of 5825 bp in length
CC   *    41940 42039: gap of     100 bp
CC   *    42040    48122: contig of 6083 bp in length
CC   *    48123 48222: gap of     100 bp
CC   *    48223    52983: contig of 4761 bp in length
CC   *    52984 53083: gap of     100 bp
CC   *    53084    58094: contig of 5011 bp in length
CC   *    58095 58194: gap of     100 bp
CC   *    58195    64606: contig of 6412 bp in length
CC   *    64607 64706: gap of     100 bp
CC   *    64707    71079: contig of 6373 bp in length
CC   *    71080 71179: gap of     100 bp
CC   *    71180    79886: contig of 8707 bp in length
CC   *    79887 79986: gap of     100 bp
CC   *    79987    86155: contig of 6169 bp in length
CC   *    86156 86255: gap of     100 bp
CC   *    86256    95370: contig of 9115 bp in length
CC   *    95371 95470: gap of     100 bp
CC   *    95471   103712: contig of 8242 bp in length
CC   *   103713 103812: gap of     100 bp
CC   *   103813   113565: contig of 9753 bp in length
CC   *   113566 113665: gap of     100 bp
CC   *   113666   123867: contig of 10202 bp in length
CC   *   123868 123967: gap of     100 bp
CC   *   123968   132296: contig of 8329 bp in length
CC   *   132297 132396: gap of     100 bp
CC   *   132397   141110: contig of 8714 bp in length
CC   *   141111 141210: gap of     100 bp
CC   *   141211   161622: contig of 20412 bp in length
CC   *   161623 161722: gap of     100 bp
CC   *   161723   189811: contig of 28089 bp in length.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .189811
FT                   /chromosome="4"
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /map="4"
FT                   /clone="RP11-392K14"
FT                   /clone_lib="RPCI-11 Human Male BAC"
FT   misc_feature    1. .451
FT                   /note="assembly_fragment clone_end:T7 vector_side:right"
FT   misc_feature    552. .2663
FT                   /note="assembly_fragment"
FT   misc_feature    2764. .4124
FT                   /note="assembly_fragment"
FT   misc_feature    4225. .6767
FT                   /note="assembly_fragment"
FT   misc_feature    6868. .8750
FT                   /note="assembly_fragment"
FT   misc_feature    8851. .11272
FT                   /note="assembly_fragment"
FT   misc_feature    11373. .13411
FT                   /note="assembly_fragment"
FT   misc_feature    13512. .16962
FT                   /note="assembly_fragment"
FT   misc_feature    17063. .22544
FT                   /note="assembly_fragment"
FT   misc_feature    22645. .26284
FT                   /note="assembly_fragment"
FT   misc_feature    26385. .31223
FT                   /note="assembly_fragment"
FT   misc_feature    31324. .36014
FT                   /note="assembly_fragment"
FT   misc_feature    36115. .41939
FT                   /note="assembly_fragment"
FT   misc_feature    42040. .48122
FT                   /note="assembly_fragment"
FT   misc_feature    48223. .52983
FT                   /note="assembly_fragment clone_end:SP6 vector_side:left"
FT   misc_feature    53084. .58094
FT                   /note="assembly_fragment"
FT   misc_feature    58195. .64606
FT                   /note="assembly_fragment"
FT   misc_feature    64707. .71079
FT                   /note="assembly_fragment"
FT   misc_feature    71180. .79886
FT                   /note="assembly_fragment"
FT   misc_feature    79987. .86155
FT                   /note="assembly_fragment"
FT   misc_feature    86256. .95370
FT                   /note="assembly_fragment"
FT   misc_feature    95471. .103712
FT                   /note="assembly_fragment"
FT   misc_feature    103813. .113565
FT                   /note="assembly_fragment"
FT   misc_feature    113666. .123867
FT                   /note="assembly_fragment"
FT   misc_feature    123968. .132296
FT                   /note="assembly_fragment"
FT   misc_feature    132397. .141110
FT                   /note="assembly_fragment"
FT   misc_feature    141211. .161622
FT                   /note="assembly_fragment"
FT   misc_feature    161723. .189811
FT                   /note="assembly_fragment"
XX
SQ   Sequence 189811 BP; 52455 A; 40839 C; 40321 G; 53485 T; 2711 other;


  Query Match             13.1%;  Score 465.8;  DB 30;  Length 189811;
  Best Local Similarity   99.6%;  Pred. No. 5.2e-91;
  Matches  467;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    416 aagtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagc 475
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97313 AGGTGCCTGAGATCATCAGCTCCATCCGTCAGGCGGGGAAGATCGCCCGGCAGGAGGAGC 97254

Qy    476 tgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcg 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97253 TGCACTGCCCGTCCGAGTTCGACGACACGTTTTCCAAGAAGTTCGAGGTGCTCTTCTGCG 97194

Qy    536 gccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgaga 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97193 GCCGCGTGACGGTGGCGCACAAGAAGGCTCCGCCGGCCCTGATCGACGAGTGCATCGAGA 97134

Qy    596 agttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatg 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97133 AGTTCAATCACGTCAGCGGCAGCCGGGGGTCCGAGAGCCCCCGCCCCAACCCGCCCCATG 97074

Qy    656 ccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagc 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97073 CCGCGCCCACAGGGAGCCAGGAGCCTGTGCGCAGGCCCATGCGCAAGTCCTTCTCCCAGC 97014

Qy    716 ccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcg 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97013 CCGGCCTGCGCTCGCTGGCCTTTAGGAAGGAGCTGCAGGATGGGGGCCTCCGAAGCAGCG 96954

Qy    776 gcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaata 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96953 GCTTCTTCAGCTCCTTCGAGGAGAGCGACATTGAGAACCACCTCATTAGCGGACACAATA 96894

Qy    836 ttgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgat 884
          |||||||||||||||||||||||||||||||||||||||||||||||| |
Db  96893 TTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGGT 96845


RESULT  11
AC021106
LOCUS       AC021106              195108 bp    DNA     linear   PRI 03-JUL-2001
DEFINITION  Homo sapiens clone RP11-177C12, complete sequence.
```

```
ACCESSION   AC021106
VERSION     AC021106.6  GI:14589687
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 195108)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 195108)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 195108)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 3, 2001 this sequence version replaced gi:14018126.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information -----------------
            Center project name: H_NH0177C12
            ------------------------------------------------------
FEATURES             Location/Qualifiers
     source          1. .195108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-177C12"
BASE COUNT    53372 a  40112 c  41873 g  59751 t
ORIGIN
```

```
  Query Match           13.1%;  Score 465.8;  DB 9;  Length 195108;
  Best Local Similarity  99.6%;  Pred. No. 5.2e-91;
  Matches  467;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    416 aagtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagc 475
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180759 AGGTGCCTGAGATCATCAGCTCCATCCGTCAGGCGGGGAAGATCGCCCGGCAGGAGGAGC 180818

Qy    476 tgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcg 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180819 TGCACTGCCCGTCCGAGTTCGACGACACGTTTTCCAAGAAGTTCGAGGTGCTCTTCTGCG 180878

Qy    536 gccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgaga 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180879 GCCGCGTGACGGTGGCGCACAAGAAGGCTCCGCCGGCCCTGATCGACGAGTGCATCGAGA 180938

Qy    596 agttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatg 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180939 AGTTCAATCACGTCAGCGGCAGCCGGGGGTCCGAGAGCCCCCGCCCCAACCCGCCCCATG 180998

Qy    656 ccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagc 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180999 CCGCGCCCACAGGGAGCCAGGAGCCTGTGCGCAGGCCCATGCGCAAGTCCTTCTCCCAGC 181058

Qy    716 ccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcg 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181059 CCGGCCTGCGCTCGCTGGCCTTTAGGAAGGAGCTGCAGGATGGGGGCCTCCGAAGCAGCG 181118

Qy    776 gcttcttcagctccttccgaggagagcgacattgagaaccacctcattagcggacacaata 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181119 GCTTCTTCAGCTCCTTCGAGGAGAGCGACATTGAGAACCACCTCATTAGCGGACACAATA 181178
```

```
Qy    836 ttgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgat 884
          |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db 181179 TTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGGT 181227
```

```
RESULT  12
AC108933/c
LOCUS       AC108933              208318 bp    DNA     linear   HTG 01-FEB-2002
DEFINITION  Homo sapiens chromosome 4 clone RP11-392K14, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AC108933 AC009595
VERSION     AC108933.1  GI:18464316
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 208318)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 208318)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 1, 2002 this sequence version replaced gi:8671945.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ---------------- Project Information -----------------
            Center project name: H_NH0392K14
            Drafting center: WIBR
            ------------------------------------------------------
            ------------- Summary Statistics  --------------
            Sequencing vector: M13; %
            Sequencing vector: plasmid; %
            Chemistry: Dye-primer ET; % of reads
            Chemistry: Dye-terminator Big Dye; % of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 205652 bases at least Q40
            Consensus quality: 206608 bases at least Q30
            Consensus quality: 207022 bases at least Q20
            Insert size: 218000; agarose-fp
            Insert size: 207918; sum-of-contigs
            Quality coverage: 6.45 in Q20 bases; agarose-fp
            Quality coverage: 6.62 in Q20 bases; sum-of-contigs
            ------------------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    25746: contig of 25746 bp in length
            *    25747    25846: gap of unknown length
            *    25847    65131: contig of 39285 bp in length
            *    65132    65231: gap of unknown length
            *    65232   105028: contig of 39797 bp in length
            *   105029   105128: gap of unknown length
            *   105129   159307: contig of 54179 bp in length
            *   159308   159407: gap of unknown length
            *   159408   208318: contig of 48911 bp in length.
FEATURES             Location/Qualifiers
     source          1. .208318
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="4"
                     /clone="RP11-392K14"
     misc_feature    1. .25746
                     /note="assembly_name:Contig16"
     misc_feature    25847. .65131
                     /note="assembly_name:Contig17"
     misc_feature    65232. .105028
                     /note="assembly_name:Contig18"
     misc_feature    105129. .159307
                     /note="assembly_name:Contig19"
     misc_feature    159408. .208318
                     /note="assembly_name:Contig20"
BASE COUNT    59195 a  45373 c  44728 g  58621 t    401 others
ORIGIN


  Query Match             13.1%;  Score 465.8;  DB 2;  Length 208318;
  Best Local Similarity   99.6%;  Pred. No. 5.2e-91;
  Matches  467;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     416 aagtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagc 475
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129553 AGGTGCCTGAGATCATCAGCTCCATCCGTCAGGCGGGGAAGATCGCCCGGCAGGAGGAGC 129494

Qy     476 tgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcg 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129493 TGCACTGCCCGTCCGAGTTCGACGACACGTTTTCCAAGAAGTTCGAGGTGCTCTTCTGCG 129434

Qy     536 gccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgaga 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129433 GCCGCGTGACGGTGGCGCACAAGAAGGCTCCGCCGGCCCTGATCGACGAGTGCATCGAGA 129374

Qy     596 agttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatg 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129373 AGTTCAATCACGTCAGCGGCAGCCGGGGGTCCGAGAGCCCCCGCCCCAACCCGCCCCATG 129314

Qy     656 ccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagc 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129313 CCGCGCCCACAGGGAGCCAGGAGCCTGTGCGCAGGCCCATGCGCAAGTCCTTCTCCCAGC 129254

Qy     716 ccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcg 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129253 CCGGCCTGCGCTCGCTGGCCTTTAGGAAGGAGCTGCAGGATGGGGGCCTCCGAAGCAGCG 129194

Qy     776 gcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaata 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129193 GCTTCTTCAGCTCCTTCGAGGAGAGCGACATTGAGAACCACCTCATTAGCGGACACAATA 129134

Qy     836 ttgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgat 884
           |||||||||||||||||||||||||||||||||||||||||||||||| |
Db  129133 TTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGGT 129085


RESULT  13
AL390202_04/c
WPCOMMENT
  Sequence split into 10 fragments  LOCUS AL390202 Accession AL390202
     Fragment Name      Begin     End
     AL390202_00            1   110000
     AL390202_01       100001   210000
     AL390202_02       200001   310000
     AL390202_03       300001   410000
     AL390202_04       400001   510000
     AL390202_05       500001   610000
     AL390202_06       600001   710000
     AL390202_07       700001   810000
     AL390202_08       800001   910000
     AL390202_09       900001   988176
Continuation (5 of 10) of AL390202 from base 400001 (AL390202 Homo sapiens chromosome 20


  Query Match             11.9%;  Score 424.2;  DB 2;  Length 110000;
  Best Local Similarity   96.0%;  Pred. No. 6.1e-82;

  Matches  435;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10365 ATGGAACCAATAACATTCACAGCAAGGAAACATCTGCTTCCTAACGAGGTCTCGGTGGAT 10306

Qy      61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10305 TTTGGCCTGCAGCTGGTGGGCTCCCTGCCTGTGCATTCCCTGACCACCATGCCCATGCTG 10246

Qy     121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10245 CCCTGGGTTGTGGCTGAGGTGCGAAGACTCAGCAGGCAGTCCACCAGAAAGGAACCTGTA 10186

Qy     181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10185 ACCAAGCAAGTCCGGCTTTGCGTTTCACCCTCTGGACTGAGATGTGAACCTGAGCCAGGG 10126

Qy     241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10125 AGAAGTCAACAGTGGGATCCCCTGATCTATTCCAGCATCTTTGAGTGCAAGCCTCAGCGT 10066

Qy     301 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10065 GTTCACAAACTGATTCACAACAGTCATGACCCAAGTTACTTTGCTTGTCTGATTAAGGAA 10006

Qy     361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaaagtg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10005 GACGCTGTCCACCGGCAGAGTATCTGCTATGTGTTCAAAGCCGATGATCAAACAAAAGTA 9946

Qy     421 cctgagatcatcagctccatccgtcaggcgggg 453
             ||||||    | |   |    | ||| | ||
Db    9945 AGTGAGATGGAGATCCAAAAGACTAAGGTGTGG 9913


RESULT  14
DMU50542
LOCUS       DMU50542               3214 bp    mRNA    linear   INV 12-JUN-1996
DEFINITION  Drosophila melanogaster pollux (plx) mRNA, complete cds.
ACCESSION   U50542
VERSION     U50542.1  GI:1373162
KEYWORDS    .
SOURCE      fruit fly strain=Oregon R.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 3214)
  AUTHORS   Zhang,S.D., Kassis,J., Olde,B., Mellerick,D.M. and Odenwald,W.F.
  TITLE     Pollux, a novel Drosophila adhesion molecule, belongs to a family
            of proteins expressed in plants, yeast, nematodes, and man
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3214)
  AUTHORS   Zhang,S.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-1996) Shang-Ding Zhang, Neurogenetics, LNC,
            NINDS, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES             Location/Qualifiers
     source          1. .3214
                     /organism="Drosophila melanogaster"
                     /strain="Oregon R"
                     /db_xref="taxon:7227"
                     /chromosome="3"
                     /map="83c"
                     /tissue_type="trachea; CNS"
     gene            1. .3214
                     /gene="plx"
     CDS             130. .2328
                     /gene="plx"
                     /function="adhesion molecule"
                     /codon_start=1
                     /product="pollux"
                     /protein_id="AAB02200.1"
```

```
                     /db_xref="GI:1373163"
                     /translation="MRKPAKRGKRDAAELRELWRTAIRQTIMLNRMETENAMLQARQN
                     ENELKRIKLDYEEIVPCDKQLIERWEQIIERNSTQIGNKKDPKVLGHAIRTGVPRSKR
                     GDVWTFLAEQHSMNTAPVDTKRFPNFNTPYHMLLKHLTEHQHAIFIDLGRTFPNHQFY
                     KDPLGLGQLSLFNLLKAYSILDPELGYCQGLGFICGVLLLHCDEANSFQLLKHLMFRR
                     NMRTKYLPDMKKFQLQLYQLSRLVKDHLPDLYVWLDQNDVSPTLYAAPWILTVFSSQF
                     PLGFVARVFDLLFLESSDVIFKFAIALLSVHKQQLLAKDNFEEIMDYLKTVVPKMEHT
                     CMEQIMKLVFSMDIGKQLAEYNVEYNVLQEEITTTNHHLEMLNREKTQNQHLEQQLQF
                     AQSSIAQLETTRSSQQAQITTLQSQVQSLELTIQTLGRYVGQLVEHNPDLELPNEVRR
                     MLQQLDDLDRQRRKPIFTERKIGKSVSVNSHLGFPLKVLEELTERDELGSPQKQKKEK
                     TPFFEQLRQQQQQHRLNGGGQSSNVGESGSPTPPSRPNRLLDNASARTVMQVKLDELK
                     LPEHVDKFVANIKSPLEVDSGVGTPLSPPSTASNSSGGSIFSRMGYRTTPPALSPLAQ
                     RQSYGVAITTAPCPQHMEEVAPATTMAVMPQEDVEEPQPMHPLSMVGGDVNVRFKGTT
                     QLKSIRPVHHMRAIPLGGVQHPSSTEPAVRVAPVPVELAPPAATATTGRS"
BASE COUNT      862 a    798 c    774 g    780 t
ORIGIN
```

Query Match 8.6%; Score 304.6; DB 3; Length 3214;
Best Local Similarity 56.0%; Pred. No. 6.1e-56;
Matches 621; Conservative 0; Mismatches 479; Indels 9; Gaps 2;

```
Qy   2155 gagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggag 2214
          |||||  | || |||||||  |  || |||   ||    ||| |||||    | ||||||
Db    169 GAGCTGAGGGAACTGTGGCGCACTGCCATTAGGCAGACTATAATGCTGAACCGCATGGAG 228

Qy   2215 aaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctc 2274
          | ||| ||    | | | ||     | |||||      |   |||||| | || ||
Db    229 ACGGAGAACGCCATGTTGCAGGCACGGCAGAATGAGAACGAGCTAAAGCGCATTAAACTG 288

Qy   2275 gattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgctt 2334
          || || || || |||   |||||    ||  |  |||        |||||  ||||  ||
Db    289 GACTACGAGGAGATTGTACCCTGCGACAAGCAGCTAATCGAGCGATGGGAGCAGATCATT 348

Qy   2335 ---agcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgtt 2391
             |||   |  |    ||  |  |||  ||     ||    ||  |  ||  ||
Db    349 GAGCGCAACTCCACACAGATAGGCAACAAGAAGGATCCCAAAGTCCTGGGCCACGCCATT 408

Qy   2392 gggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattc 2451
            |    || || || ||     || || || ||  ||||||  || || |||||||||
Db    409 CGCACCGGAGTACCGCGATCGAAGCGCGGCGATGTCTGGACCTTCCTGGCTGAGCAGCAT 468

Qy   2452 caccttaa------acaccagtttcccagcaaacagcagccaaaggatgtgccatacaaa 2505
            | | ||      ||    |  | |    |  |          ||   || ||  |
Db    469 TCCATGAACACGGCACCGGTGGACACAAAGCGATTCCCCAACTTCAATACACCGTATCAC 528

Qy   2506 gaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacc 2565
            ||  | || || || |||  || ||||||||||| | |||| || ||   | ||
Db    529 ATGCTGCTGAAACACCTTACTGAGCATCAGCATGCGATTTTCATTGATCTGGGCAGGACT 588

Qy   2566 tttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaac 2625
          |||||||  ||||  | || | |  |  ||||||   || ||| | |||| |  |||
Db    589 TTTCCTAATCACCAGTTCTACAAGGATCCGCTTGGTCTCGGCCAGTTGTCGCTGTTTAAC 648

Qy   2626 attttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagcttt 2685
            |  ||||||| |||||  |||||| | |||| | || || |||||  |||||  ||||
Db    649 CTGCTGAAGGCGTACTCCATTCTAGATCCGGAATTAGGGTACTGCCAGGGTCTGGGCTTC 708

Qy   2686 gtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagttt 2745
            |    ||| | || || || |||     ||| |    |  | ||| || |||| |||
Db    709 ATCTGCGGCGTCTTACTCCTGCATTGCGATGAAGCCAATTCATTTCAACTGCTGAAGCAC 768

Qy   2746 ctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatc 2805
          || ||||||  |       ||||  |    | || | ||| |||||||  |  ||  ||  |
Db    769 CTAATGTTTCGCCGTAATATGCGCACGAAATACCTGCCGGACATGAAAAAGTTTCAACTG 828

Qy   2806 cagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctg 2865
          |||  |||| ||||||||  |  |||  |   | |||  | |    |||  |||||      ||
Db    829 CAGCTGTATCAGCTCTCCCGATTGGTCAAGGATCATCTGCCAGATCTCTACGTGTGGCTC 888

Qy   2866 gaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcc 2925
          ||  || ||||  ||   |||||  | || || || || ||| ||||||| | |||  |
Db    889 GATCAGAACGATGTCTCGCCCACTTTGTATGCGGCTCCATGGATCCTCACCGTCTTTAGC 948
```

```
Qy   2926 tcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaaca 2985
          |||||||| |||||||||||||| ||| | |||||||| || | || ||   |    |
Db     949 TCACAGTTTCCGCTGGGATTCGTGGCCCGCGTCTTTGATCTGCTCTTCCTGGAATCCTCC 1008

Qy   2986 gaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcag 3045
          || || || ||||| | ||  |     || |       || ||||      | |||
Db    1009 GATGTAATCTTTAAGTTTGCCATTGCCTTGCTCTCCGTGCACAAGCAGCAACTACTGGCC 1068

Qy   3046 catgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttg 3105
            | || ||| | |||   ||  | |||| | | || | |  | | || ||  | |
Db    1069 AAGGATAACTTCGAAGAGATTATGGACTATCTGAAGACCGTGGTGCCAAAGATGGAGCAC 1128

Qy   3106 gtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaa 3165
                |||||  | |  || || | ||| |||   |||||||| |  || ||| |
Db    1129 ACGTGCATGGAGCAAATTATGAAGCTGGTCTTTAGCATGGACATTGGAAAGCAGCTTGCC 1188

Qy   3166 gcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagt 3225
          |  ||  | || |||||| | || ||||| || ||  ||| |    |    | || |
Db    1189 GAATACAATGTGGAGTACAATGTGCTTCAGGAGGAGATTACCACCACTAACCATCACCTA 1248

Qy   3226 gacaaccaaagaatggataaattagagaa 3254
          || |  |  |    ||| ||     ||||
Db    1249 GAAATGCTTAACCGGGAGAAGACGCAGAA 1277
```

```
RESULT 15
DME17919
LOCUS       DME17919                4746 bp    mRNA    linear   INV 18-NOV-1998
DEFINITION  Drosophila melanogaster mRNA for pollux protein.
ACCESSION   Y17919
VERSION     Y17919.1  GI:3893102
KEYWORDS    pollux gene.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 4746)
  AUTHORS   Xu,X.Z., Wes,P.D., Chen,H., Li,H.S., Yu,M., Morgan,S., Liu,Y. and
            Montell,C.
  TITLE     Retinal targets for calmodulin include proteins implicated in
            synaptic transmission
  JOURNAL   J. Biol. Chem. 273 (47), 31297-31307 (1998)
  MEDLINE   99030403
REFERENCE   2  (bases 1 to 4746)
  AUTHORS   Montell,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-1998) C. Montell, The Johns Hopkins University
            School of Medicine, Department of Biological Chermistry, 725 N.
            Wolfe Street, Baltimore, MD 21205-2185, USA
FEATURES             Location/Qualifiers
     source          1. .4746
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
     source          1. .3316
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /map="83C"
                     /tissue_type="retinal"
                     /dev_stage="adult"
     gene            452. .4591
                     /gene="pollux"
     CDS             452. .4591
                     /gene="pollux"
                     /note="calmodulin binding protein"
                     /codon_start=1
                     /product="Pollux protein"
                     /protein_id="CAA76939.1"
                     /db_xref="GI:3893103"
                     /translation="MAELMHQMRDPAHTLGGSVGSIPQTLIGGGGGSHGNSNGALNGI
                     HATPATNLKMSEAMRNAQHDTSANPVSSKMKASKSYTHGLSSSSGTVNIPTSTSAQSN
                     LSLLADISPNHTHFFEVMYVGKIRVSQKRVPNTFIDDALPKFKAYDAQRLRLLQNRKM
```

```
                     SLSSEGGVGIEAKPSSSLKSHDLKEEDEEEQEQHKGHDDSQDSQAKPLVQLQLTGAEE
                     GAAPRPLEDNKENKSPEKRPLLRGQSQIELGHKEHSDGSQPSAANSQLEAPNVIVNKQ
                     PTPPRDQGVGTGTASASAGPSQLHPNYAMDNIPKQRDRSASQGCIPPYVEQNRTMVFL
                     VGRCDLRLISPDRKQVLLYKDFKDVASCVHGQKSLDHFGIICRELNNDGYIGYVFKCQ
                     SEHVCDDIVAAIAQAFDTCAEQKKKQDTQIFSCEHCPMLWYHKLCTDVEGLSEKKTQA
                     LILRRIETLSDDEQEIVWAKFCGSEKTNSPVAEQNQFLMMLLRAHCESRQQRHVHDTA
                     ENRSEFLNQYLGGSTIFMKAKRSLTNSFDNLLKRKPSKDDIAVPSHNLRDIREGSAEP
                     LGTQSPPEGFRSRSNTVGASPSSKPTAEQLKSPMMDIFIKVGNSPKEAETHQGSWRQA
                     ILNSVVTPSKGLDSEVPTEFLSPMRKPAKRGKRDAAELRELWRTAIRQTIMLNRMETE
                     NAMLQARQNENELKRIKLDYEEIVPCDKQLIERWEQIIERNSTQIGNKKDPKVLGHAI
                     RTGVPRSKRGDVWTFLAEQHSMNTAPVDTKRFPNFNTPYHMLLKHLTEHQHAIFIDLG
                     RTFPNHQFYKDPLGLGQLSLFNLLKAYSILDPELGYCQGLGFICGVLLLHCDEANSFQ
                     LLKHLMFRRNMRTKYLPDMKKFQLQLYQLSRLVKDHLPDLYVWLDQNDVSPTLYAAPW
                     ILTVFSSQFPLGFVARVFDLLFLESSDVIFKFAIALLSVHKQQLLAKDNFEEIMDYLK
                     TVVPKMEHTCMEQIMKLVFSMDIGKQLAEYNVEYNVLQEEITTTNHHLEMLNREKTQN
                     QHLEQQLQFAQSSIAQLETTRSSQQAQITTLQSQVQSLELTIQTLGRYVGQLVEHNPD
                     LELPNEVRRMLQQLDDLDRQRRKPIFTERKIGKSVSVNSHLGFPLKVLEELTERDELG
                     SPQKQKKEKTPFFEQLRQQQQQHRLNGGGQSSNVGESGSPTPPSRPNRLLDNASARTV
                     MQVKLDELKLPEHVDKFVANIKSPLEVDSGVGTPLSPPSTASNSSGGSIFSRMGYRTT
                     PPALSPLAQRQSYGVAITTAPCPHHMEEVAPATTMAVMPQEDVEEPQPMHPLSMVGGD
                     VNVRFKGTTQLKSIRPVHHMRAIPLGGVQHPSSTEPAVRVAPVPVELAPPAATATTGR
                     S"
BASE COUNT     1211 a   1342 c   1261 g    932 t
ORIGIN


  Query Match             8.6%;  Score 304.6;  DB 3;  Length 4746;
  Best Local Similarity   56.0%;  Pred. No. 6.4e-56;
  Matches  621;  Conservative    0;  Mismatches  479;  Indels    9;  Gaps    2;

Qy    2155 gagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggag 2214
           |||||  | || |||||||  ||  || |||   ||    ||| |||||    | ||||||
Db    2432 GAGCTGAGGGAACTGTGGCGCACTGCCATTAGGCAGACTATAATGCTGAACCGCATGGAG 2491

Qy    2215 aaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctc 2274
           | ||| ||    | | | || ||     | |||||      |   |||||| | || ||
Db    2492 ACGGAGAACGCCATGTTGCAGGCACGGCAGAATGAGAACGAGCTAAAGCGCATTAAACTG 2551

Qy    2275 gattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgctt 2334
           || || || || |||   |||||    ||  |  |||        |||||  ||||  ||
Db    2552 GACTACGAGGAGATTGTACCCTGCGACAAGCAGCTAATCGAGCGATGGGAGCAGATCATT 2611

Qy    2335 ---agcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgtt 2391
              |||  |  |     |  |  |||   ||     ||    ||  |  ||  ||
Db    2612 GAGCGCAACTCCACACAGATAGGCAACAAGAAGGATCCCAAAGTCCTGGGCCACGCCATT 2671

Qy    2392 gggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattc 2451
            |    || || || ||      | || || ||  ||||||  || || |||||||||
Db    2672 CGCACCGGAGTACCGCGATCGAAGCGCGGCGATGTCTGGACCTTCCTGGCTGAGCAGCAT 2731

Qy    2452 caccttaa------acaccagtttcccagcaaacagcagccaaaggatgtgccatacaaa 2505
             | | ||      ||    |  | |    |   |          ||   || ||  |
Db    2732 TCCATGAACACGGCACCGGTGGACACAAAGCGATTCCCCAACTTCAATACACCGTATCAC 2791

Qy    2506 gaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacc 2565
             ||  | || || || |||  || ||||||||||||| | ||||| || ||  | ||
Db    2792 ATGCTGCTGAAACACCTTACTGAGCATCAGCATGCGATTTTCATTGATCTGGGCAGGACT 2851

Qy    2566 tttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaac 2625
           |||||||  ||||  | || |  |  | ||||||    || ||| | ||||| |  |||
Db    2852 TTTCCTAATCACCAGTTCTACAAGGATCCGCTTGGTCTCGGCCAGTTGTCGCTGTTTAAC 2911

Qy    2626 attttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagcttt 2685
            |  ||||||| |||||  |||||||| | |||| | || || |||||  |||||  ||||
Db    2912 CTGCTGAAGGCGTACTCCATTCTAGATCCGGAATTAGGGTACTGCCAGGGTCTGGGCTTC 2971

Qy    2686 gtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagttt 2745
            |   ||| | || || || |||    ||| |   |  | ||| || |||| |||
Db    2972 ATCTGCGGCGTCTTACTCCTGCATTGCGATGAAGCCAATTCATTTCAACTGCTGAAGCAC 3031

Qy    2746 ctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatc 2805
           || ||||||  |       |||| |   | || | ||| |||||||  |  || ||  |
Db    3032 CTAATGTTTCGCCGTAATATGCGCACGAAATACCTGCCGGACATGAAAAAGTTTCAACTG 3091

Qy    2806 cagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctg 2865
           ||| |||| |||||||||  | ||| |  | ||| | |    ||| ||||||      ||
Db    3092 CAGCTGTATCAGCTCTCCCGATTGGTCAAGGATCATCTGCCAGATCTCTACGTGTGGCTC 3151

Qy    2866 gaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcc 2925
           ||  || |||  ||   ||||   | || || || || ||| |||||||| | |||  |
Db    3152 GATCAGAACGATGTCTCGCCCACTTTGTATGCGGCTCCATGGATCCTCACCGTCTTTAGC 3211

Qy    2926 tcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaaca 2985
           ||||||||  |||||||||||||| ||| | |||||||||  || | || ||  |    |
Db    3212 TCACAGTTTCCGCTGGGATTCGTGGCCCGCGTCTTTGATCTGCTCTTCCTGGAATCCTCC 3271

Qy    2986 gaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcag 3045
           || || || |||||  | ||  |      || |       || ||||      | |||
Db    3272 GATGTAATCTTTAAGTTTGCCATTGCCTTGCTCTCCGTGCACAAGCAGCAACTACTGGCC 3331

Qy    3046 catgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttg 3105
            | || ||| | |||   ||  | |||| | | || | |  | | || ||  | |
Db    3332 AAGGATAACTTCGAAGAGATTATGGACTATCTGAAGACCGTGGTGCCAAAGATGGAGCAC 3391

Qy    3106 gtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaa 3165
                |||||  | |  || || | ||| |||   |||||||| |  || ||| |
Db    3392 ACGTGCATGGAGCAAATTATGAAGCTGGTCTTTAGCATGGACATTGGAAAGCAGCTTGCC 3451

Qy    3166 gcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagt 3225
           |  ||  | || |||||| | || ||||| || ||  ||| |    |    | || |
Db    3452 GAATACAATGTGGAGTACAATGTGCTTCAGGAGGAGATTACCACCACTAACCATCACCTA 3511

Qy    3226 gacaaccaaagaatggataaattagagaa 3254
           || |  |  |    ||| ||     ||||
Db    3512 GAAATGCTTAACCGGGAGAAGACGCAGAA 3540


Search completed: September  9, 2002, 19:34:43
Job time: 18596 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 14:24:52 ; Search time 129.72 Seconds
(without alignments)
6731.630 Million cell updates/sec

Title: US-09-762-311-3_COPY_171_3725
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

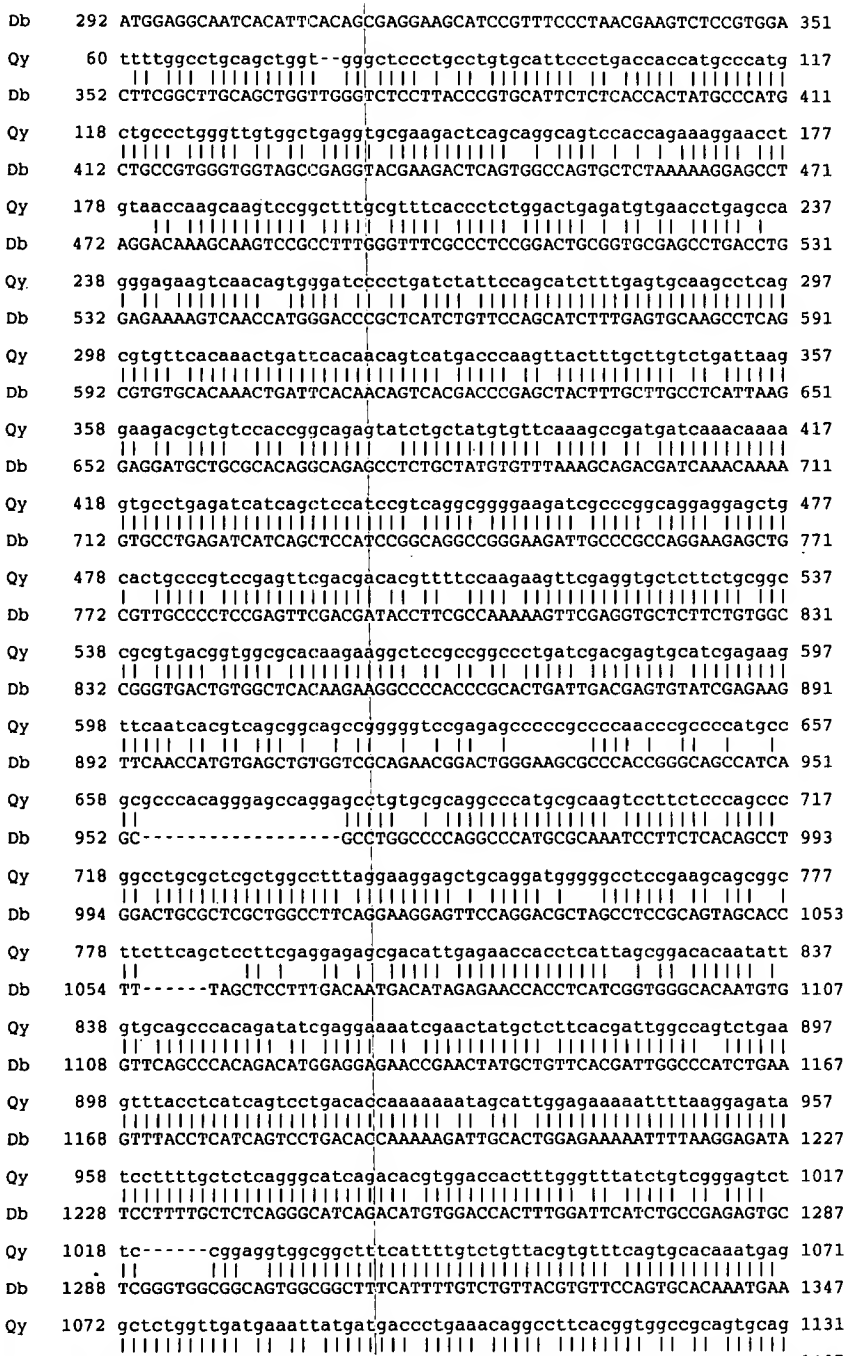
Result No.	Score	% Query Match	Length	DB	ID	Description
1	2510.4	70.6	4039	1	US-08-363-300-1	Sequence 1, Appli
c 2	57.8	1.6	7218	1	US-08-232-463-14	Sequence 14, Appl
3	45.4	1.3	9515	1	US-08-920-812-13	Sequence 13, Appl
4	45.4	1.3	9515	1	US-08-920-827-13	Sequence 13, Appl
5	45.4	1.3	9515	1	US-08-921-177-13	Sequence 13, Appl
6	45.4	1.3	9515	1	US-08-362-577C-13	Sequence 13, Appl
7	45.4	1.3	9515	2	US-08-920-828-13	Sequence 13, Appl
8	45	1.3	8201	1	US-08-253-155A-9	Sequence 9, Appli
9	41	1.2	1957	1	US-08-295-060-3	Sequence 3, Appli
10	40.8	1.1	1030	3	US-08-858-003-2	Sequence 2, Appli
11	40.8	1.1	1030	3	US-09-078-166-2	Sequence 2, Appli
12	40.8	1.1	1030	4	US-08-997-467-2	Sequence 2, Appli
13	40.6	1.1	2249	3	US-08-814-052-19	Sequence 19, Appl
14	40.6	1.1	2279	3	US-08-814-052-17	Sequence 17, Appl
15	40.6	1.1	2300	3	US-08-814-052-18	Sequence 18, Appl
16	40.6	1.1	3183	2	US-08-939-218A-1	Sequence 1, Appli
17	40.6	1.1	3192	1	US-08-940-661A-1	Sequence 1, Appli
18	40.6	1.1	3192	2	US-09-083-485-1	Sequence 1, Appli
c 19	40.6	1.1	3358	3	US-09-248-571-2	Sequence 2, Appli
20	40.4	1.1	4134	4	US-09-162-021B-1	Sequence 1, Appli
21	40	1.1	1962	2	US-08-919-421A-1	Sequence 1, Appli
22	40	1.1	1962	4	US-09-342-819-1	Sequence 1, Appli
23	39.2	1.1	3187	5	PCT-US95-06815-1	Sequence 1, Appli
24	39.2	1.1	3192	1	US-08-706-037-26	Sequence 26, Appl
25	39.2	1.1	3192	2	US-09-005-397-26	Sequence 26, Appl
26	39	1.1	2241	2	US-08-838-219B-20	Sequence 20, Appl
27	39	1.1	2241	3	US-09-233-336A-20	Sequence 20, Appl
28	39	1.1	2241	3	US-09-233-752A-20	Sequence 20, Appl
29	39	1.1	2241	4	US-09-402-036-20	Sequence 20, Appl
30	39	1.1	2370	2	US-08-838-219B-19	Sequence 19, Appl
31	39	1.1	2370	3	US-09-233-336A-19	Sequence 19, Appl
32	39	1.1	2370	3	US-09-233-752A-19	Sequence 19, Appl
33	39	1.1	2370	4	US-09-402-036-19	Sequence 19, Appl
34	39	1.1	2403	1	US-08-471-033-30	Sequence 30, Appl
35	39	1.1	2403	2	US-08-471-044-30	Sequence 30, Appl
36	39	1.1	2403	2	US-08-463-483A-30	Sequence 30, Appl
37	39	1.1	2403	2	US-08-471-046A-30	Sequence 30, Appl
38	39	1.1	2403	2	US-08-470-566B-30	Sequence 30, Appl
39	39	1.1	2403	2	US-08-838-219B-7	Sequence 7, Appli
40	39	1.1	2403	2	US-08-469-334-30	Sequence 30, Appl
41	39	1.1	2403	3	US-09-300-529-30	Sequence 30, Appl
42	39	1.1	2403	3	US-09-233-336A-7	Sequence 7, Appli
43	39	1.1	2403	3	US-09-233-752A-7	Sequence 7, Appli
44	39	1.1	2403	4	US-09-402-036-7	Sequence 7, Appli
c 45	38.8	1.1	1069	4	US-09-372-422A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-363-300-1
; Sequence 1, Application US/08363300
; Patent No. 5700927
; GENERAL INFORMATION:
; APPLICANT: Zon, Leonard and Richardson, Paul
; TITLE OF INVENTION: Tbc1 Gene and Uses Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,300
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04590/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4039 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..3829
US-08-363-300-1

Query Match 70.6%; Score 2510.4; DB 1; Length 4039;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 2949; Conservative 0; Mismatches 536; Indels 37; Gaps 8;

Qy 1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59


Db 292 ATGGAGGCAATCACATTCACAGCGAGGAAGCATCCGTTTCCCTAACGAAGTCTCCGTGGA 351
Qy 60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
Db 352 CTTCGGCTTGCAGCTGGTTGGGTCTCCTTACCCGTGCATTCTCTCACCACTATGCCCATG 411
Qy 118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
Db 412 CTGCCGTGGGTGGTAGCCGAGGTACGAAGACTCAGTGGCCAGTGCTCTAAAAAGGAGCCT 471
Qy 178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
Db 472 AGGACAAAGCAAGTCCGCCTTTGGGTTTCGCCCTCCGGACTGCGGTGCGAGCCTGACCTG 531
Qy 238 gggagaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcag 297
Db 532 GAGAAAAGTCAACCATGGGACCCGCTCATCTGTTCCAGCATCTTTGAGTGCAAGCCTCAG 591
Qy 298 cgtgttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
Db 592 CGTGTGCACAAACTGATTCACAACAGTCACGACCCGAGCTACTTTGCTTGCCTCATTAAG 651
Qy 358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
Db 652 GAGGATGCTGCGCACAGGCAGAGCCTCTGCTATGTGTTTAAAGCAGACGATCAAACAAAA 711
Qy 418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
Db 712 GTGCCTGAGATCATCAGCTCCATCCGGCAGGCCGGGAAGATTGCCCGCCAGGAAGAGCTG 771
Qy 478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
Db 772 CGTTGCCCCTCCGAGTTCGACGATACCTTCGCCAAAAAGTTCGAGGTGCTCTTCTGTGGC 831
Qy 538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
Db 832 CGGGTGACTGTGGCTCACAAGAAGGCCCCACCCGCACTGATTGACGAGTGTATCGAGAAG 891
Qy 598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
Db 892 TTCAACCATGTGAGCTGTGGTCGCAGAACGGACTGGGAAGCGCCCACCGGGCAGCCATCA 951
Qy 658 gcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagccc 717
Db 952 GC----------GCCTGGCCCCAGGCCCATGCGCAAATCCTTCTCACAGCCT 993
Qy 718 ggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggc 777
Db 994 GGACTGCGCTCGCTGGCCTTCAGGAAGGAGTTCCAGGACGCTAGCCTCCGCAGTAGCACC 1053
Qy 778 ttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatatt 837
Db 1054 TT------TAGCTCCTTTGACAATGACATAGAGAACCACCTCATCGGTGGGCACAATGTG 1107
Qy 838 gtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaa 897
Db 1108 GTTCAGCCCACAGACATGGAGGAGAACCGAACTATGCTGTTCACGATTGGCCCATCTGAA 1167
Qy 898 gtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaattttaaggagata 957
Db 1168 GTTTACCTCATCAGTCCTGACACCAAAAAGATTGCACTGGAGAAAAATTTTAAGGAGATA 1227
Qy 958 tccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtct 1017
Db 1228 TCCTTTTGCTCTCAGGGCATCAGACATGTGGACCACTTTGGATTCATCTGCCGAGAGTGC 1287
Qy 1018 tc------cggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
Db 1288 TCGGGTGGCGGCAGTGGCGGCTTTCATTTTGTCTGTTACGTGTTCCAGTGCACAAATGAA 1347
Qy 1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
Db 1348 GCTCTGGTTGACGAGATCATGATGACTCTGAAGCAGGCTTTCACGGTAGCTGCGGTGCAG 1407
Qy 1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
Db 1408 CAGACGGCTAAGGCACCAGCCCAGCTCTGTGAGGGCTGCCCCTTGCAAGGCCTGCACAAG 1467
Qy 1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
Db 1468 CTCTGCGAAAGGATAGAGGGAATGAATTCATCTAAAACCAAATTAGAACTCCAGAAGCAC 1527
Qy 1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
Db 1528 TTGACCACACTGACCAATCAGGAGCAGGCCACCATATTCGAGGAGGTTCAGAAATTGAGA 1587
Qy 1312 ccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaa 1371
Db 1588 CCAAGAAACGAGCAGCGAGAGAATGAATTAATTATTTCTTTTCTGAGGTGCTTATATGAA 1647
Qy 1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
Db 1648 GAGAAGCAAAAAGAGCACAGCCACACTGGGGCGCCAAAGCAGACACTACAGGTGGCAGCA 1707
Qy 1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
Db 1708 GAGAATATTGGGAGTGACCTGCCACCCAGTGCTAGCCGGTTCAGGTTAGATTCGCTGAAG 1767
Qy 1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1551
Db 1768 AACAGAGCAAAGAGGTCCTTAACAGAGTCCCTAGAGAGCATTCTGTCCCGGGGTAATAAA 1827
Qy 1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
Db 1828 GCCAGAGGCCTGCAGGACCATTCCGCCAGTGTGGATCTGGACAGCTCCACTTCTAGTACT 1887
Qy 1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
Db 1888 CTAAGTAACACCAGCAAAGAGCTGTCCATGGGTGACAAGGAGGCCTTCCCCGTCTCTGAG 1947
Qy 1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
Db 1948 ACCTCCTTCAAGCTCCTTGGCTCCTCAGATGACCTGTCCAGTGACTCAGAGGGCCACATT 2007
Qy 1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
Db 2008 GCAGAAGAGTCTGCCCTGTTGTCACCCCAGCAGGCGTTCAGAAGGAGAGCCAACACCCTG 2067
Qy 1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
Db 2068 AGTCATTTCCCAGTAGAGTGCCCTGCGCCTCCAGAACCTGCCCAGAGCTCTCCAGGGGTC 2127
Qy 1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
Db 2128 TCTCAAAGGAAACTCATGCGGTACCACTCCGTGAGCACAGAGACGCCTCATGAACGCAAG 2187
Qy 1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
Db 2188 GACTTTGAATCCAAAGCAAACCACCTGGGTGACACAGATGGGACCCCCGTGAAGACCCGG 2247
Qy 1972 aggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgat 2031
Db 2248 CGGCACTCGTGGAGACAGCAGATATTCCTTCGAGTGGCCACTCCACAGAAGGCTTGTGAC 2307
Qy 2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
Db 2308 TCCCCGAGCAGATATGAAGATTATTCCGAGCTGGGAGAGCTCCCTCCACGCTCCCCTTTA 2367
Qy 2092 gaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaaaggacatct 2151
Db 2368 GAACCGGTGTGTGAGGACGGCCCATTTGG-CCAGTACAGGAAGAAAAGAGGAAGACGTCA 2426
Qy 2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg-ctgcttagaat 2210
Db 2427 CGCGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTGCAGCAGATCCTGCCTCGTCAGGAT 2486

```
Qy   2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
          |||||||| |||||||||| |||||||||||||| |||||||||||| ||||| ||
Db   2487 GGAGAAGGAGAATCAGAAGCTACAAGCCTCTGAAAACGATTTGCTGAACAAACGCCTCAA 2546

Qy   2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
          ||| || |||||||||||| ||||| |||||||||||||| ||||||||||||||||||
Db   2547 GCTTGACTATGAAGAAATCACTCCGTGTCTTAAAGAAGTCACTACAGTGTGGGAAAAGAT 2606

Qy   2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||||
Db   2607 GCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAAGTGCACTCAGCTGT 2666

Qy   2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||| ||
Db   2667 TGGGCAAGGTGTGCCACGTCATCACCGAGGTGAGATCTGGAAATTTCTAGCTGAGCAGTT 2726

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
          |||||||||||||| ||||| || |||||||||||||||| ||||| |||||||| ||
Db   2727 CCACCTTAAACACCCATTTCCTAGTAAACAGCAGCCAAAGGACGTGCCCTACAAAGAGCT 2786

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaaccttttcc 2570
          | | ||| ||||||| || |||||||| || ||||| || ||||| ||||||||||||||
Db   2787 CCTGAAGAAGCTGACCTCGCAGCAGCACGCCATTCTCATCGACCTCGGGCGAACCTTTCC 2846

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
          ||||| |||||||||||||||||||||||||||||| ||||| || |||||||||||| |
Db   2847 AACACATCCATACTTCTCTGCCCAGCTTGGAGCAGGTCAGCTGTCACTTTACAACATTCT 2906

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
          |||||||||||| ||||| |||||||| || ||||| |||||||||||||||||||| ||
Db   2907 GAAGGCCTACTCGCTTCTGGACCAGGAGGTTGGATACTGCCAAGGTCTCAGCTTTGTGGC 2966

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
          |||||||||||||||||| |||||||||||||||||| || |||||||||| ||||| |
Db   2967 AGGCATTTTGCTTCTTCACATGAGTGAGGAAGAGGCGTTCAAGATGCTCAAGTTCCTGAT 3026

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3027 GTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTGCAGATCCAGAT 3086

Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          |||||||| ||  || | || || |||||||| |||||||||||| |||||||| ||
Db   3087 GTACCAGCTGTCACGGCTCCTCCACGATTACCACCGAGACCTCTACAACCACCTGGAAGA 3146

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          ||||||||  ||||||      |||||| || |||||||| |||||| |||| ||||||||
Db   3147 GCACGAGACTGGCCCCCCTACGTACGCGGCTCCCTGGTTTCTCACCGTGTTCGCCTCACA 3206

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          |||||| || || || |||||||||||||||||||||| || ||||||||| |||||||
Db   3207 GTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGATCAGAGGT 3266

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          |||||||||||| ||||| |||||||||||||||||||||||| |||||||||||||||
Db   3267 CATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTACAGCATGA 3326

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
          ||||| |||||||| || |||||| ||||| | ||| || |||||||| ||| |||| ||
Db   3327 GAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGCCTGGTGCA 3386

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          |||||| |||||||||| |||||| ||||| |||||||||| || ||| | || || ||
Db   3387 GATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTCCAGGCCTA 3446

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||| || |||||||||||| | || || || ||||| || ||||| ||||||||||||||
Db   3447 TGAGGTCGAGTACCACGTCGTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTCAGTGACAA 3506

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          |||||||||||| ||||| |||||||||||||||| || ||||||||||||||||||||
Db   3507 CCAAAGAATGGAGAAATTGGAGAAAACCAACAGCACGTTGCGCAAACAGAACCTTGACCT 3566

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
          ||| || |||||||||||||||||| |||||||||||||| |||||  | |||||
Db   3567 CCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACGGTAGAGAA 3626

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
          || || | ||| ||||| |||||||||||     ||| ||| | ||  |||||||
Db   3627 ACTTCTTACCAGCGAGAGTAAGCTGAAGCAGCGTGCGCTGACCCTGGAGGTGGAGCG--T 3684

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
          |||||||||||||| |||||||||||||||| |||  |||||   |||||||   |  ||
Db   3685 CGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGCACTCCAGA 3744

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
          ||| || ||||| |||| |||||||| ||||| |||| |||
Db   3745 GCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 3786
```

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 1.6%; Score 57.8; DB 1; Length 7218;
Best Local Similarity 2.9%; Pred. No. 9.7e-06;
Matches 11; Conservative 225; Mismatches 147; Indels 0; Gaps 0;

```
Qy   1198 gagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacg 1257
          ::::::: :::::::: :::       :::: ::::  ::::  : ::::: :   :: :
Db   1428 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1369
```

```
Qy   1258 acattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgaga 1317
Db   1368 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1309

Qy   1318 aatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaa 1377
Db   1308 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1249

Qy   1378 cagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaat 1437
Db   1248 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1189

Qy   1438 attggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaa 1497
Db   1188 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1129

Qy   1498 gcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccaga 1557
Db   1128 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1069

Qy   1558 ggcctgcaggaacactccatcag 1580
Db   1068 RRATCGCAAGCTCCCTCGACCTG 1046


RESULT   3
US-08-920-812-13
; Sequence 13, Application US/08920812
; Patent No. 5763188
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,812
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-920-812-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagcccccgccccaacccgccccatgccgcgcccacagggagccagga 677
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   4
US-08-920-827-13
; Sequence 13, Application US/08920827
; Patent No. 5770375
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,827
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-920-827-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy     438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
           |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db     802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy     498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
           || |  || | |  | |   ||||   ||||| |  |     || | |||  || ||  |
Db     862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy     558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
           ||      |  |          ||    |      | ||     |       ||     | |
Db     922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy     618 ccgggggtccgagagccccccgcccaaccgccccatgccgcgcccacagggagccagga 677
           |   |  |||  |   ||  |   || |  |||  |||  | || | || | |  ||
Db     982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy     678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
           |    |      | ||  ||   | | ||||||   |   |||||  |  ||||||| |
Db    1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy     738 taggaaggagctgcaggatgggggcct 764
           |  || | ||| ||| ||||  ||
Db    1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   5
US-08-921-177-13
; Sequence 13, Application US/08921177
; Patent No. 5798211
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/921,177
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-921-177-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy     438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
           |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db     802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy     498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
           || |  || | |  | |   ||||   ||||| |  |     || | |||  || ||  |
Db     862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy     558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
           ||      |  |          ||    |      | ||     |       ||     | |
Db     922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy     618 ccgggggtccgagagccccccgcccaaccgccccatgccgcgcccacagggagccagga 677
           |   |  |||  |   ||  |   || |  |||  |||  | || | || | |  ||
Db     982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy     678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
           |    |      | ||  ||   | | ||||||   |   |||||  |  ||||||| |
Db    1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy     738 taggaaggagctgcaggatgggggcct 764
           |  || | ||| ||| || |  ||
Db    1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   6
US-08-362-577C-13
; Sequence 13, Application US/08362577C
; Patent No. 5807673
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/362,577C
;      FILING DATE:  27-MAR-1995
```

```
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-362-577C-13


  Query Match             1.3%;  Score 45.4;  DB 1;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
          |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
          || |  || | |  | |   ||||   ||||| |  |     || | ||| || ||  |
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
          ||       |  |          ||    |       | ||    |      ||     | |
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagcccccgccccaacccgccccatgccgcgcccacagggagccagga 677
          |    |  ||| |  ||  | | || |  ||| ||| | || | || | | ||
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
          |   |     | ||  ||   | | ||||||   |   |||||  |  |||||||  |
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
            |  || | ||| ||| || |  ||
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   7
US-08-920-828-13
; Sequence 13, Application US/08920828
; Patent No. 5853998
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,828
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Pseudomonas aeruginosa
;      STRAIN:  Clinical Isolate P2-2
US-08-920-828-13


  Query Match             1.3%;  Score 45.4;  DB 2;  Length 9515;
  Best Local Similarity   46.2%;  Pred. No. 0.033;
  Matches  151;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy    438 catccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcga 497
          |||  |  || ||| || |  |||| || |    ||  | | | || |  || |  ||
Db    802 CATGGGCAAGCCGGTGATGGACGCCTGGAACATCGATGTACCCGGCGCCGCCCACGTCTT 861

Qy    498 cgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaa 557
          || |  || | |  | |   ||||   ||||| |  |     || | ||| || ||  |
Db    862 CGCCTGGTATGCGGAAAGCCTCGACAAGCTCTACGACCAGGTCGCGCCGGCCGCCCAGCA 921

Qy    558 gaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcag 617
          ||       |  |          ||    |       | ||    |      ||     | |
Db    922 GACCCTGGCCACCATTACCCGCGTGCCGCTGGGGGTGATCGGCGCGGTGGTGCCGTGGAA 981

Qy    618 ccgggggtccgagagcccccgccccaacccgccccatgccgcgcccacagggagccagga 677
          |    |  ||| |  ||  | | || |  ||| ||| | || | || | | ||
Db    982 CTTCCCGCTCGACATGGCCGCCTGGAAGCTCGCCCCGGCCCTGGCCGCCGGCAACTCGGT 1041

Qy    678 gcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctgcgctcgctggcctt 737
          |   |     | ||  ||   | | ||||||   |   |||||  |  |||||||  |
Db   1042 GGTGCTCAAGCCGGCCGAGCAGTCGCCGTTCTCCGCCCTGCGCCTGGCCGAGCTGGCCCT 1101

Qy    738 taggaaggagctgcaggatgggggcct 764
            |  || | ||| ||| || |  ||
Db   1102 GGAGGCGGGGGTGCCGGAAGGCGTGCT 1128


RESULT   8
US-08-253-155A-9
; Sequence 9, Application US/08253155A
; Patent No. 5691147
;  GENERAL INFORMATION:
;    APPLICANT:  Gyuris, Jeno
;    APPLICANT:  Draetta, Giulio
;    TITLE OF INVENTION:  CDK4 Binding Proteins
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
```

```
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII(text)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/253,155A
;       FILING DATE:  02-JUN-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Vincent, Matthew P.
;       REGISTRATION NUMBER:  36,709
;       REFERENCE/DOCKET NUMBER:  MII-028
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 227-7400
;       TELEFAX:  (617) 227-5941
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  8201 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-253-155A-9


  Query Match             1.3%;  Score 45;  DB 1;  Length 8201;
  Best Local Similarity   53.0%;  Pred. No. 0.039;
  Matches   96;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy    2549 ttgaccttgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggac 2608
           | ||| |  ||  ||  | | |  ||    | |||    |  |    ||||||   |
Db    2133 TGGACGTGAGGACGACTCTCCGGAACCATGTCTTCTTTAGGGATCGATATGGAGCCAAGC 2192

Qy    2609 agctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatatt 2668
           ||     || | | ||||  |   |||||| ||        ||| ||| ||||| || |
Db    2193 AGAGGGAACTATTCTACATCCTCCTGGCCTATTCGGAGTATAACCCGGAGGTGGGCTACT 2252

Qy    2669 gccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgt 2728
           ||   |  || |||    |  | | | | || || ||| || ||  |||||| || || |
Db    2253 GCAGGGACCTGAGCCACATCACCGCCTTGTTCCTCCTTTATCTGCCTGAGGAGGACGCAT 2312

Qy    2729 t 2729
           |
Db    2313 T 2313


RESULT   9
US-08-295-060-3
; Sequence 3, Application US/08295060
; Patent No. 5659123
;  GENERAL INFORMATION:
;    APPLICANT:  VAN RIE, Jeroen
;    APPLICANT:  JANSENS, Stefan
;    APPLICANT:  PERFEROEN, Marnix
;    TITLE OF INVENTION:  NEW DIABROTICA TOXINS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/295,060
;      FILING DATE:  26-AUG-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  010830-052
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1957 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  13..1947
US-08-295-060-3


  Query Match             1.2%;  Score 41;  DB 1;  Length 1957;
  Best Local Similarity   47.9%;  Pred. No. 0.22;
  Matches  113;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy     488 ccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacgg 547
           || ||  | ||||||       |||   |  ||  |||  ||    | ||  ||    |
Db    1322 CCCAGACCTACGACAGCNNNCGCAACGTGGGCGCTGTGAGCTGGGACAGCATCGACCAGC 1381

Qy     548 tggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacg 607
           || | | |||    |      | |||||    | |  | || | |  || | ||  |||
Db    1382 TGCCTCCCGAGACCACCGACGAGCCCCTGGAGAAGGGCTACAGCCACCAGCTGAACTACG 1441

Qy     608 tcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcgcccacag 667
           | |   ||  || |  |   |  |||    || ||| |||    |   ||  | ||
Db    1442 TGATGTGCTTCCTGATGCAGGGCAGCAGGGGCACCATCCCAGTGCTGACCTGGACCCACA 1501

Qy     668 ggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggcctg 723
           ||||  |||     |   || |  |    |||||    || ||||||    ||||
Db    1502 AGAGCGTGGACTTCTTCAACATGATCGACAGCAAGAAGATCACCCAGCTGCCCCTG 1557


RESULT  10
US-08-858-003-2
; Sequence 2, Application US/08858003
; Patent No. 6060234
;  GENERAL INFORMATION:
;    APPLICANT:  Katz, Leonard
;    APPLICANT:  Stassi, Diane L.
;    APPLICANT:  Summers Jr., Richard G.
;    APPLICANT:  Ruan, Xiaoan
;    APPLICANT:  Pereda-Lopez, Ana
;    APPLICANT:  Kakavas, Stephan J.
;    TITLE OF INVENTION:  NOVEL POLYKETIDE DERIVATIVES
;    TITLE OF INVENTION:  AND RECOMBINANT METHODS FOR MAKING SAME
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Rd.
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/858,003
;      FILING DATE:  16-MAY-1979
```

```
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne Casuto
;      REGISTRATION NUMBER:  P-40,943
;      REFERENCE/DOCKET NUMBER:  4952.US.P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (847)-938-3137
;      TELEFAX:  (847)-938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-858-003-2


  Query Match             1.1%;  Score 40.8;  DB 3;  Length 1030;
  Best Local Similarity   50.0%;  Pred. No. 0.18;
  Matches  102;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy    440 tccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcgacg 499
          ||   |||||||    | |||||  | ||      |||  | | | || ||||  | |||
Db    203 TCACACAGGCGGCTCTGTTCGCCGTGGAGACCTCCCTGTTCCGGCTGTTCGAGGCCCACG 262

Qy    500 acacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaaga 559
          |    |  || |  |  ||  | | || |  ||||   ||||| | ||| |||  |
Db    263 GCCTCGTCCCCGACTACCTCATCGGCCACTCCATCGGCGAAGTGACCGCGGCCCACCTGG 322

Qy    560 aggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcagcc 619
            |     |  |  |||   |||| |||| ||  |     | | | || | ||   |  |
Db    323 CCGGGGTCCTCGATCTGGCGGACGCGTGCGTCCTGGTCGCCCACCGCGGCCGCCTGATGC 382

Qy    620 gggggtccgagagcccccgcccca 643
            |  | ||  |  |  | || |||
Db    383 AGTCGGCCCGGGCCGGCGGCGCGA 406


RESULT  11
US-09-078-166-2
; Sequence 2, Application US/09078166
; Patent No. 6063561
;  GENERAL INFORMATION:
;    APPLICANT:  Katz, Leonard
;    APPLICANT:  Stassi, Diane L.
;    APPLICANT:  Summers Jr., Richard G.
;    APPLICANT:  Ruan, Xiaoan
;    APPLICANT:  Pereda-Lopez, Ana
;    APPLICANT:  Kakavas, Stephan J.
;    TITLE OF INVENTION:  NOVEL POLYKETIDE DERIVATIVES
;    TITLE OF INVENTION:  AND RECOMBINANT METHODS FOR MAKING SAME
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Rd.
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/078,166
;      FILING DATE:  16-MAY-1979
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne Casuto
;      REGISTRATION NUMBER:  P-40,943
;      REFERENCE/DOCKET NUMBER:  4952.US.P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (847)-938-3137
;      TELEFAX:  (847)-938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-078-166-2


  Query Match             1.1%;  Score 40.8;  DB 3;  Length 1030;
  Best Local Similarity   50.0%;  Pred. No. 0.18;
  Matches  102;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy    440 tccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcgacg 499
          ||   |||||||    | |||||  | ||      |||  | | | || ||||  | |||
Db    203 TCACACAGGCGGCTCTGTTCGCCGTGGAGACCTCCCTGTTCCGGCTGTTCGAGGCCCACG 262

Qy    500 acacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaaga 559
          |    |  || |  |  ||  | | || |  ||||   ||||| | ||| |||  |
Db    263 GCCTCGTCCCCGACTACCTCATCGGCCACTCCATCGGCGAAGTGACCGCGGCCCACCTGG 322

Qy    560 aggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcagcc 619
            |     |  |  |||   |||| |||| ||  |     | | | || | ||   |  |
Db    323 CCGGGGTCCTCGATCTGGCGGACGCGTGCGTCCTGGTCGCCCACCGCGGCCGCCTGATGC 382

Qy    620 gggggtccgagagcccccgcccca 643
            |  | ||  |  |  | || |||
Db    383 AGTCGGCCCGGGCCGGCGGCGCGA 406


RESULT  12
US-08-997-467-2
; Sequence 2, Application US/08997467
; Patent No. 6200813
;  GENERAL INFORMATION:
;    APPLICANT:  Katz, Leonard
;    APPLICANT:  Stassi, Diane L.
;    APPLICANT:  Summers Jr., Richard G.
;    APPLICANT:  Ruan, Xiaoan
;    APPLICANT:  Pereda-Lopez, Ana
;    APPLICANT:  Kakavas, Stephan J.
;    TITLE OF INVENTION:  NOVEL POLYKETIDE DERIVATIVES
;    TITLE OF INVENTION:  AND RECOMBINANT METHODS FOR MAKING SAME
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Rd.
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,467
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/858,003
;      FILING DATE:  16-MAY-1997
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne Casuto
;      REGISTRATION NUMBER:  P-40,943
;      REFERENCE/DOCKET NUMBER:  4952.US.P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (847)-938-3137
;      TELEFAX:  (847)-938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-997-467-2


  Query Match             1.1%;  Score 40.8;  DB 4;  Length 1030;
  Best Local Similarity   50.0%;  Pred. No. 0.18;
  Matches  102;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy    440 tccgtcaggcggggaagatcgcccggcaggaggagctgcactgcccgtccgagttcgacg 499
          ||   |||||||    | |||||  | ||      |||  | | | || ||||  | |||
Db    203 TCACACAGGCGGCTCTGTTCGCCGTGGAGACCTCCCTGTTCCGGCTGTTCGAGGCCCACG 262

Qy    500 acacgttttccaagaagttcgaggtgctcttctgcggccgcgtgacggtggcgcacaaga 559
          |   |  || |  |  ||  |  | || |  ||||   ||||| | ||| |||  |
Db    263 GCCTCGTCCCCGACTACCTCATCGGCCACTCCATCGGCGAAGTGACCGCGGCCCCACCTGG 322

Qy    560 aggctccgccggccctgatcgacgagtgcatcgagaagttcaatcacgtcagcggcagcc 619
            |     |  |  |||   |||| |||| ||  |     | | | || | ||   |  |
Db    323 CCGGGGTCCTCGATCTGGCGGACGCGTGCGTCCTGGTCGCCCACCGCGGCCGCCTGATGC 382

Qy    620 gggggtccgagagccccccgcccca 643
            |  | ||  |  |  | || | |
Db    383 AGTCGGCCCGGGCCGGCGGCGCGA 406


RESULT  13
US-08-814-052-19
; Sequence 19, Application US/08814052
; Patent No. 6015783
;  GENERAL INFORMATION:
;    APPLICANT:  von der Osten, Claus
;    APPLICANT:  Cherry, Joel R.
;    APPLICANT:  Bjornvad, Mads E.
;    APPLICANT:  Vind, Jesper
;    APPLICANT:  Rasmussen, Michael Dolberg
;    TITLE OF INVENTION:  PROCESS FOR REMOVAL OR BLEACHING OF SOILING
;    TITLE OF INVENTION:  OR STAINS FROM CELLULOSIC FABRIC
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6015783o No. 6015783disk of No. 6015783th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,052
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4684.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-814-052-19


  Query Match             1.1%;  Score 40.6;  DB 3;  Length 2249;
  Best Local Similarity   48.9%;  Pred. No. 0.31;
  Matches  109;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy    474 gctgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctg 533
          |||   | ||   |||  |     | || ||    |||       | | |    |  |
Db   1234 GCTCTGCGGCGGCTCCAGGAATCCCTACCCGGCCGCCATCTTCCACTACGCCGGCGCCCC 1293

Qy    534 cggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcga 593
          |||| ||  | |   ||   |   ||||| |||   | ||  | |  |  |  | ||
Db   1294 CGGCGGCCCGCCCACGGACGAGGGCAAGGCCCCGGTCGACCACAACTGCCTGGACCTCCC 1353

Qy    594 gaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgcccca 653
            ||  |||| | ||||   | | ||   | | ||  ||||   |  | |||| | |||| |
Db   1354 CAACCTCAAGCCCGTCGTGGCCCGCGACGTGCCCCTGAGCGGCTTCGCCAAGCGGCCCGA 1413

Qy    654 tgccgcgcccacagggagccaggagcctgtgcgcaggcccatg 696
            | ||| |   |  | ||  ||  |   | ||| |||| ||
Db   1414 CAACACGCTCGACGTCACCCTCGACACCACGGGCACGCCCCTG 1456


RESULT  14
US-08-814-052-17
; Sequence 17, Application US/08814052
; Patent No. 6015783
;  GENERAL INFORMATION:
;    APPLICANT:  von der Osten, Claus
;    APPLICANT:  Cherry, Joel R.
;    APPLICANT:  Bjornvad, Mads E.
;    APPLICANT:  Vind, Jesper
;    APPLICANT:  Rasmussen, Michael Dolberg
;    TITLE OF INVENTION:  PROCESS FOR REMOVAL OR BLEACHING OF SOILING
;    TITLE OF INVENTION:  OR STAINS FROM CELLULOSIC FABRIC
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6015783o No. 6015783disk of No. 6015783th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,052
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4684.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2279 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-814-052-17


  Query Match             1.1%;  Score 40.6;  DB 3;  Length 2279;
  Best Local Similarity   48.9%;  Pred. No. 0.32;
  Matches  109;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy     474 gctgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctg 533
           |||   | ||   |||  |    | || ||    |||       | | |    |  |
Db    1340 GCTCTGCGGCGGCTCCAGGAATCCCTACCCGGCCGCCATCTTCCACTACGCCGGCGCCCC 1399

Qy     534 cggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcga 593
           ||||  ||  | |   ||   |   ||||| |||   | ||  | |  |  |  |  | ||
Db    1400 CGGCGGCCCGCCCACGGACGAGGGCAAGGCCCCGGTCGACCACAACTGCCTGGACCTCCC 1459

Qy     594 gaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgcccca 653
            ||  |||| | ||||   | | ||   | | ||  ||||  |  | |||| | |||| |
Db    1460 CAACCTCAAGCCCGTCGTGGCCCGCGACGTGCCCCTGAGCGGCTTCGCCAAGCGGCCCGA 1519

Qy     654 tgccgcgcccacagggagccaggagcctgtgcgcaggcccatg 696
             | ||| |   |  | ||  ||  |   | ||| |||| ||
Db    1520 CAACACGCTCGACGTCACCCTCGACACCACGGGCACGCCCCTG 1562


RESULT  15
US-08-814-052-18
; Sequence 18, Application US/08814052
; Patent No. 6015783
;  GENERAL INFORMATION:
;    APPLICANT:  von der Osten, Claus
;    APPLICANT:  Cherry, Joel R.
;    APPLICANT:  Bjornvad, Mads E.
;    APPLICANT:  Vind, Jesper
;    APPLICANT:  Rasmussen, Michael Dolberg
;    TITLE OF INVENTION:  PROCESS FOR REMOVAL OR BLEACHING OF SOILING
;    TITLE OF INVENTION:  OR STAINS FROM CELLULOSIC FABRIC
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6015783o No. 6015783disk of No. 6015783th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,052
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4684.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2300 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-814-052-18


  Query Match             1.1%;  Score 40.6;  DB 3;  Length 2300;
  Best Local Similarity   48.9%;  Pred. No. 0.32;
  Matches  109;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy     474 gctgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctg 533
           |||   | ||   |||  |    | || ||    |||       | | |    |  |
Db    1234 GCTCTGCGGCGGCTCCAGGAATCCCTACCCGGCCGCCATCTTCCACTACGCCGGCGCCCC 1293

Qy     534 cggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcga 593
           ||||  ||  | |   ||   |   ||||| |||   | ||  | |  |  |  |  | ||
Db    1294 CGGCGGCCCGCCCACGGACGAGGGCAAGGCCCCGGTCGACCACAACTGCCTGGACCTCCC 1353

Qy     594 gaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgcccca 653
            ||  |||| | ||||   | | ||   | | ||  ||||  |  | |||| | |||| |
Db    1354 CAACCTCAAGCCCGTCGTGGCCCGCGACGTGCCCCTGAGCGGCTTCGCCAAGCGGCCCGA 1413

Qy     654 tgccgcgcccacagggagccaggagcctgtgcgcaggcccatg 696
            | ||| |   |  | ||  ||  |   | ||| |||| ||
Db    1414 CAACACGCTCGACGTCACCCTCGACACCACGGGCACGCCCCTG 1456


Search completed:  September  9, 2002, 19:06:22
Job time:  16890 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 14:07:12 ; Search time 5191.84 Seconds
(without alignments)
9241.748 Million cell updates/sec

Title: US-09-762-311-3_COPY_171_3725
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	833.6	23.4	1001	10	Z78359	Z78359 HSZ78359 Hu
	2	826.6	23.3	1051	10	BM479578	BM479578 AGENCOURT
	3	819.2	23.0	1007	10	BM016445	BM016445 603641252
	4	782.4	22.0	784	10	BG828092	BG828092 602753463
	5	741.8	20.9	835	10	BG759286	BG759286 602710975
	6	738.8	20.8	800	10	BG746377	BG746377 602703644
	7	737.4	20.7	1121	10	BM476629	BM476629 AGENCOURT
	8	729.8	20.5	1109	10	BM460573	BM460573 AGENCOURT
	9	722.2	20.3	952	10	BE299948	BE299948 600944494
	10	713.8	20.1	838	10	BI907513	BI907513 603065640
	11	705	19.8	807	10	BG120498	BG120498 602346815
	12	701	19.7	832	10	BM014940	BM014940 603640877
	13	689.8	19.4	851	9	AL570425	AL570425 AL570425
	14	683.4	19.2	754	10	BE300005	BE300005 600944594
	15	682.6	19.2	892	10	BE561889	BE561889 601346094
	16	678.6	19.1	686	10	BG769758	BG769758 602744658
	17	675	19.0	1108	10	BG326728	BG326728 602425604
	18	671.8	18.9	978	10	BG029927	BG029927 602297271
	19	669.6	18.8	670	9	AL557623	AL557623 AL557623
c	20	669.4	18.8	857	9	AL544177	AL544177 AL544177
	21	668.4	18.8	750	10	BI084159	BI084159 602869636
	22	664.6	18.7	795	10	BE560184	BE560184 601346896
	23	661	18.6	661	9	AL046787	AL046787 DKFZp434M
	24	661	18.6	948	10	BG258116	BG258116 602379212
	25	659.2	18.5	672	10	BE675153	BE675153 7f03d07.x
	26	652.8	18.4	806	9	AW043925	AW043925 wy82b07.x
	27	648.8	18.3	665	9	AW173375	AW173375 xj78g07.x
	28	647.8	18.2	812	10	BI758489	BI758489 603023334
	29	641	18.0	1161	10	BM466841	BM466841 AGENCOURT
	30	637.8	17.9	773	10	BE561312	BE561312 601344450
c	31	636.4	17.9	862	10	BI910369	BI910369 603068025
	32	635.4	17.9	660	10	BE276812	BE276812 601178843
	33	632.4	17.8	761	10	BF797592	BF797592 602257390
	34	632	17.8	753	10	BF305442	BF305442 601892937
	35	630	17.7	722	10	BE560392	BE560392 601347108
	36	628	17.7	748	10	BG424359	BG424359 602447963
	37	624.4	17.6	627	9	AW772402	AW772402 hn73a08.x
	38	621.2	17.5	626	9	AW239183	AW239183 xb37a10.y
	39	612	17.2	623	10	BE884440	BE884440 601510724
	40	604.4	17.0	649	10	BE513375	BE513375 601315702
	41	597	16.8	605	9	AW732976	AW732976 bb18b11.y
	42	596.2	16.8	622	10	BE514108	BE514108 601315945
	43	592.4	16.7	774	10	BI763670	BI763670 603049558
	44	590.2	16.6	952	10	BG424757	BG424757 602453405
	45	589.4	16.6	758	10	BG166636	BG166636 602339021

ALIGNMENTS

RESULT 1
Z78359/c
LOCUS Z78359 1001 bp mRNA linear EST 28-JUL-1999
DEFINITION HSZ78359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone 3.114 (CEPH) 5', mRNA sequence.
ACCESSION Z78359
VERSION Z78359.1 GI:1495132
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1001)
AUTHORS Neri,C., Albanese,V., Lebre,A.S., Holbert,S., Saada,C., Bougueleret,L., Meier-Ewert,S., LeGall,I., Millasseau,P., Bui,H., Giudicelli,C., Massart,C., Guillou,S., Gervy,P., Poullier,E., Rigault,P., Weissenbach,J., Lennon,G., Chumakov,I., Dausset,J., Lehrach,H., Cohen,D. and Cann,H.M.
TITLE Survey of CAG/CTG repeats in human cDNAs representing new genes: candidates for inherited neurological disorders
JOURNAL Hum. Mol. Genet. 5 (7), 1001-1009 (1996)
MEDLINE 96414310
COMMENT Contact: Neri C.
Fondation Jean Dausset - CEPH
27 Rue Juliette Dodu, 75010 Paris, France
Related sequence: Z78360
5'-sequence (upper strand).
FEATURES Location/Qualifiers
source 1. .1001
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3.114 (CEPH)"
/clone_lib="Human fetal brain S. Meier-Ewert"
/tissue_type="brain"
/dev_stage="fetus"
/note="cDNA library of S. Meier-Ewert, Max Planck Inst.f.Mol.Genetics, Berlin, FRG"
BASE COUNT 237 a 229 c 231 g 272 t 32 others
ORIGIN

```
  Query Match           23.4%;  Score 833.6;  DB 10;  Length 1001;
  Best Local Similarity  93.5%;  Pred. No. 2.2e-197;
  Matches  937;  Conservative    0;  Mismatches   56;  Indels    9;  Gaps    8;

Qy   2146 acatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgctt 2205
          ||||||||||||||||||||||||| | ||||||||||||||||||||| |||||||
Db   1001 ACATCTCGTGAGCTCCGAGAGCTGTGCCTAAAGGCTATTCTTCAACAGATNCTGCTGCTN 942

Qy   2206 agaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgc 2265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    941 AGAATGGAGAAGGAAAATCAGAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGC 882

Qy   2266 ctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaa 2325
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    881 CTGAAGCTCGATTATGAAGAAATNACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAA 822

Qy   2326 aagatgcttagcactccaggaagat--caaaaattaagtttgacatggaaaaaatgcact 2383
          ||||||||||||||||||||||||  || ||||||||||||||||||||||||||||||
Db    821 AAGATGCTTAGCACTCCAGGAAGATCCCANAAATTAAGTTTGACATGGAAAAAATGCACT 762

Qy   2384 cggctgttgggc-aaggtgtgccacgtcatcaccgaggtgaaatctggaaa-tttctagc 2441
          || |||||||| |||||||||||||||||||||||||||||||||||||| ||||||||
Db    761 CGACTGTTGGGCTAAGGTGTGCCACGTCATCACCGAGGTGAAATCTGGAAATTTTCTAGC 702

Qy   2442 tgagcaa-ttccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccat 2500
          ||||||| ||||||| |||||||||||||||||| |||||||||| ||||||||||||
Db    701 TGAGCAATTTCCACCGTAAACACCAGTTTCCCAGCANACAGCAGCCANAGGATGTGCCAT 642

Qy   2501 acaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggc 2560
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||| |
Db    641 ACAAAGAACTCTTNAAGCAGCTGACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGNC 582

Qy   2561 gaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgcttt 2620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    581 GAACCTTTCCTACACACCCATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTT 522

Qy   2621 acaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctca 2680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ACAACATTTTGAAGGCCTACTCACTTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCA 462

Qy   2681 gctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctca 2740
          |||||||| |||||||||| |||||||||||||||||||||||||| ||| |||||||
Db    461 GCTTTGTACCAGGCATTTTGCGTCTTCATATGAGTGAGGAAGAGGCGTGTAAGATGCTCA 402

Qy   2741 agtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttac 2800
          |||||||||| |||||||||||  ||  |||| |||||||||||||||||||||||||||
Db    401 AGTTTCTGATG-TTGACATGGGCCTNNGGGNACAGNATCGGCCAGACATGATTATTTTAC 343

Qy   2801 agatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatc 2860
          ||| |||||||||| |||| | | |||||||||||||||||||||||||||||||||||
Db    342 AGANCCAGATGTACCANCTCTCGGGGTTNCNTNATGATTACCACAGAGACCTCAACAATC 283

Qy   2861 acctggaggagcacgagatcggccccagcctctacgctg-ccccctggttcctcaccatg 2919
          |||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||
Db    282 ACCTGGAGGAGGNCGAGATCGGCCCCAGCCTCTACGCTGCCCCCCCTGGTTCCTCACCATG 223

Qy   2920 tttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcag 2979
          |||||||||||||| |||||||||||||  ||||| |||||| ||||||||||||||||
Db    222 TTTGCCTCACAGTTGCCGCTGGGATTCGTATNCAGAGNCTTTGANATGATTTTTCTTCAG 163

Qy   2980 ggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgatt 3039
          |||||||||||||||||||||||||   ||| ||| |||||| |||||||||  |||||
Db    162 GGAACAGAGGTCATATTTAAAGTGGCGGNAAGNCTGNTGGGAANCCATAAGCCNGTGATT 103

Qy   3040 ctgcagcatgaaaacctagaaa-ccatagttgactttataaaaagcacgctacccaacct 3098
          ||||| ||||||||||||||  ||| ||| |||||||| ||||||||||||||||||||
Db    102 CTGCANCATGAAAACCTAGAANGCCAGAGTNGACTTTATGTAAAGCACGCTACCCAACCT 43

Qy   3099 tggcttggtacagatggaa-aagaccatcaatcaggtatttg 3139
          |||| ||||  |||||||  ||||||||||||||||  ||
Db     42 TGGCGTGGTNNAGATGGAAGNAGACCATCAATCAGGTAGGTG 1
```

```
RESULT   2
BM479578
LOCUS       BM479578               1051 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6464998 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5577183
            5', mRNA sequence.
ACCESSION   BM479578
VERSION     BM479578.1  GI:18528620
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1051)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12330  row: d  column: 16
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .1051
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5577183"
                     /clone_lib="NIH_MGC_92"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      296 a    278 c    234 g    243 t
ORIGIN
```

```
  Query Match           23.3%;  Score 826.6;  DB 10;  Length 1051;
  Best Local Similarity  95.8%;  Pred. No. 1.3e-195;
  Matches  892;  Conservative    0;  Mismatches   34;  Indels    5;  Gaps    4;

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCACCTTAAACACCAGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACT 60

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTAAAGCAGCTGACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCC 120

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TACACACCCATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTT 180

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAAGGCCTACTCACTTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGC 240

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AGGCATTTTGCTTCTTCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGAT 300

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGAT 360
```

```
Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTACCAGCTCTCGAGGTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGA 420

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCACGAGATCGGCCCCAGCCTCTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACA 480

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GTTCCCGCTGGGATTCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGT 540

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CATATTTAAAGTGGCTTTAAGTCTGTTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGA 600

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAACCTAGAAACCATAGTTGACTTTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACA 660

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GATGGAAAAGACCATCAATCAGGTATTTGAAATGGACATCGCTAAACAGTTACAAGCTTA 720

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGAAGTTGAGTACCACGTCCTTCAAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAA 780

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          || ||||||||||| |     ||| | |       ||||||||||||  ||||||||||
Db    781 CC-AAGAATGGATAATTAGAGAAAACCAACAGCAGGCTTACGCAAACGAAACCTTGACCC 839

Qy   3291 ccttgaacagttgcaggt-ggcaaatggtaggatccaaagccttg--aggccaccattga 3347
          | |  |||||||||||| ||| ||||||| ||||||||||| ||    | || |||| ||
Db    840 TCCTTGACAGTTGCAGGTGGGCCAATGGTAAGATCCAAAGCCCTGGAGGCCCCCCATGGA 899

Qy   3348 gaagct-cctgagcagtgagagcaagctgaa 3377
          |||||| |||||| ||| |||||||||||  ||
Db    900 AAAGCTCCCTGAACAGGGAGACCAAGCCTAA 930


RESULT   3
BM016445
LOCUS       BM016445               1007 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603641252F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5417148 5',
            mRNA sequence.
ACCESSION   BM016445
VERSION     BM016445.1  GI:16530799
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1007)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12064  row: h  column: 13
            High quality sequence stop: 822.
FEATURES             Location/Qualifiers
     source          1. .1007
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5417148"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      274 a    246 c    255 g    231 t      1 others
ORIGIN


  Query Match             23.0%;  Score 819.2;  DB 10;  Length 1007;
  Best Local Similarity   95.1%;  Pred. No. 8.9e-194;
  Matches  910;  Conservative    0;  Mismatches   39;  Indels    8;  Gaps    6;

Qy   2532 gcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatacttctctgc 2591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGC 60

Qy   2592 ccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcacttctaga 2651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAGA 120

Qy   2652 ccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttcttcatat 2711
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    121 CCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCANTTTGCTTCTTCATAT 180

Qy   2712 gagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaa 2771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAA 240

Qy   2772 acagtatcggccagacatgattatttttacagatccagatgtaccagctctcgaggttgct 2831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACAGTATCGGCCAGACATGATTATTTTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCT 300

Qy   2832 tcatgattaccacagagacctctacaatcacctggaggagcacgagatcggccccagcct 2891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCT 360

Qy   2892 ctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagc 2951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGC 420

Qy   2952 cagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtggctttaag 3011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAG 480

Qy   3012 tctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaaccatagttga 3071
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCTGTCGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGA 540

Qy   3072 ctttataaaaagcacgctacccaaccttggcttggtacagatggaaaagaccatcaatca 3131
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTTTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCA 600

Qy   3132 ggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtaccacgtcct 3191
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGTATCTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCT 660

Qy   3192 tcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataa-attag 3250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db    661 TCAAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAACCTTAG 720

Qy   3251 agaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgca-ggtg 3309
          |||||||||||||||| ||||||||||||||||  |||||||||||||||||| |||| 
Db    721 AGAAAACCAACAGCAG-TTACGCAAACAGAACCTTGGACTCCTTGAACAGTTGCAGGGTG 779

Qy   3310 gcaaatggtaggatccaaagccttgaggccaccattgagaa--gctcctgagcagtgaga 3367
          |||||||||||||||||||||||||||||||| ||||||||     || ||||||||||
Db    780 GCAAATGGTAGGATCCAAAGCCTTGAGGCCCCCATTGAGAAAGCTCCCGGAGCAGTGAGA 839

Qy   3368 gcaag--ctgaagcaggccatgcttacc-ttagaactggagcggtcggccctgctgcaga 3424
```

```
          |||||    |||||||||||||||||| ||||||||||    |   | ||||||||||||
Db    840 GCAAGGCTGGAAGCAGGCCATGCTTACCTTTAGAACTGGGAGCGGTCGGCCTGCTGCAGA 899

Qy   3425 cggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgca 3481
          |||||||||| ||||||  | ||| |    |  ||      |||||||| |||||||
Db    900 CGGTGGAGGAACTGCGGGCGGGGAACCGCAAAGCCCAGCGACGGGAGCCGGAGTGCA 956


RESULT   4
BG828092
LOCUS       BG828092                 784 bp    mRNA    linear   EST 22-MAY-2001
DEFINITION  602753463F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4906191 5',
            mRNA sequence.
ACCESSION   BG828092
VERSION     BG828092.1  GI:14175679
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1808  row: n  column: 16
            High quality sequence stop: 784.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4906191"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      236 a    190 c    182 g    176 t
ORIGIN


  Query Match             22.0%;  Score 782.4;  DB 10;  Length 784;
  Best Local Similarity   99.9%;  Pred. No. 1.3e-184;
  Matches  783;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1859 ggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttg 1918
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGAAACTTATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTG 60

Qy   1919 aatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcatt 1978
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCAAAGCAAACCATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATT 120

Qy   1979 cctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttcca 2038
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCTGGAGGCAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCA 180

Qy   2039 gcagatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccag 2098
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    181 GCAGATATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAG 240

Qy   2099 tttgtgaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagc 2158
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTTGTGAAGATGGGCCCTTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGC 300

Qy   2159 tccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaagg 2218
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCCGAGAGCTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGG 360

Qy   2219 aaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgatt 2278
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAATCAGAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATT 420

Qy   2279 atgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagca 2338
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATGAAGAAATTACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCA 480

Qy   2339 ctccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaag 2398
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCCAGGAAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAG 540

Qy   2399 gtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttta 2458
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGTGCCACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTA 600

Qy   2459 aacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagc 2518
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AACACCAGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGC 660

Qy   2519 agctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacc 2578
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    661 AGCTGACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACAAACC 720

Qy   2579 catacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcct 2638
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCT 780

Qy   2639 actc 2642
           ||||
Db    781 ACTC 784


RESULT   5
BG759286
LOCUS       BG759286                 835 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602710975F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851090 5',
            mRNA sequence.
ACCESSION   BG759286
VERSION     BG759286.1  GI:14069939
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 835)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1693  row: f  column: 19
            High quality sequence stop: 766.
FEATURES             Location/Qualifiers
     source          1. .835
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4851090"
                     /clone_lib="NIH_MGC_48"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      220 a    230 c    214 g    171 t
ORIGIN


  Query Match             20.9%;  Score 741.8;  DB 10;  Length 835;
  Best Local Similarity   98.8%;  Pred. No. 1.9e-174;
  Matches  768;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy   1591 gatagctccctgtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaag 1650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCATCTGTGTGTGAAAAG 60

Qy   1651 gaggccttgcccatctctgagagctcctttaagctcctcggctcctcggaggacctgtcc 1710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCCTCGGAGGACCTGTCC 120

Qy   1711 agtgactcggagagtcatctcccagaagagccagctccgctgtcgccccagcaggccttc 1770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCGCCCCAGCAGGCCTTC 180

Qy   1771 aggaggcgagcaaacaccctgagtcacttccccatcgaatgccaggaacctccacaacct 1830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAGGAACCTCCACAACCT 240

Qy   1831 gcccgggggtccccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcaca 1890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTATCACTCAGTGAGCACA 300

Qy   1891 gagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggt 1950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTTGGTGATTCTGGT 360

Qy   1951 gggactcctgtgaagacccggaggcattcctggaggcagcagatattcctccgagtagcc 2010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATTCCTCCGAGTAGCC 420

Qy   2011 accccgcagaaggcgtgcgattcttccagcagatatgaagattattcagagctgggagag 2070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTCAGAGCTGGGAGAG 480

Qy   2071 cttcccccacgatctcctttagaaccagtttgtgaagatgggcccttggccccccacca 2130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCCTTTGGCCCCCCACCA 540

Qy   2131 gaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaa 2190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAAGGCTATTCTTCAA 600

Qy   2191 cagatactgctgcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgat 2250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGCCTCTGAACATGAT 660

Qy   2251 ttgctgaacaagcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagta 2310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTGCTGAACAAGCGCCTGAAGCTCGATTATGACGAAATTACTCCCTGTCTTAAAGAAGTA 720

Qy   2311 actacagtgt-gggaaaagatgctta-gcactccaggaagatcaaaaattaagtttg 2365
           |||||||||| ||||| |||||||| |||||||||||||||||| ||| | ||| |
Db    721 ACTACAGTGTGGGGAACAGATGCTTAGGCACTCCAGGAAGATCAACAATCAGGTTCG 777
```

```
RESULT   6
BG746377
LOCUS       BG746377                 800 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602703644F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4856891 5',
            mRNA sequence.
ACCESSION   BG746377
VERSION     BG746377.1  GI:14057030
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 800)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1708  row: h  column: 12
            High quality sequence stop: 762.
FEATURES             Location/Qualifiers
     source          1. .800
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4856891"
                     /clone_lib="NIH_MGC_15"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      213 a    218 c    203 g    166 t
ORIGIN


  Query Match             20.8%;  Score 738.8;  DB 10;  Length 800;
  Best Local Similarity   98.6%;  Pred. No. 1.1e-173;
  Matches  787;  Conservative    0;  Mismatches    7;  Indels    4;  Gaps    4;

Qy   1502 agagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcc 1561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 AGAGATCTTTAACAGAGTCTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCC 61

Qy   1562 tgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaaca 1621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 TGCAGGAACACTCCATCAGTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACA 121

Qy   1622 ccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttta 1681
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 CCAGCAAAGAGCCATCTGTGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTA 181

Qy   1682 agctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagc 1741
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 AGCTCCTCGGCTCCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGC 241

Qy   1742 cagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcc 1801
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 CAGCTCCGCTGTCGCCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCC 301
```

```
Qy   1802 ccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagga 1861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 CCATCGAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGA 361

Qy   1862 aacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaat 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 AACTTATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAAT 421

Qy   1922 ccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcct 1981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 CCAAAGCAAACCATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCT 481

Qy   1982 ggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagca 2041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 GGAGGCAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCA 541

Qy   2042 g-atatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2100
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 GCATATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTT 601

Qy   2101 tgtgaa-gatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagct 2159
          |||||| ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    602 TGTGAACGATGGGCCCTTTGGCCCCCCACCAGAGGAACAGAAAAGGACATCTCGTGAGCT 661

Qy   2160 ccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatgga-gaagg 2218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    662 CCGAGAGCTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGACGAAGG 721

Qy   2219 aaaatcagaagctccaagcctctgaaaatgatttgctg-aacaagcgcctgaagctcgat 2277
          |||||||||||||||| |||||| |||||||||||| ||| ||||||||||||||||||
Db    722 CAAATCAGAAGCTCCAAGCTTCTGAACATGATTTGCTGAAACCAGCGCCTGAAGCTCGAT 781

Qy   2278 tatgaagaaattactccc 2295
          |||||||||| ||||||
Db    782 CATGAAGAAATCACTCCC 799
```

```
RESULT   7
BM476629
LOCUS       BM476629                1121 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6476337 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5553589
            5', mRNA sequence.
ACCESSION   BM476629
VERSION     BM476629.1  GI:18525671
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1121)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12270  row: m  column: 14
            High quality sequence start: 6
            High quality sequence stop: 636.
FEATURES             Location/Qualifiers
     source          1. .1121
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5553589"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      312 a    277 c    280 g    251 t      1 others
ORIGIN
```

```
  Query Match             20.7%;  Score 737.4;  DB 10;  Length 1121;
  Best Local Similarity   96.1%;  Pred. No. 2.8e-173;
  Matches  799;  Conservative    0;  Mismatches   26;  Indels    6;  Gaps    4;

Qy    830 acaatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggcc 889
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     31 ACTATATTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGATTGGCC 90

Qy    890 agtctgaagtttacctcatcagtcctgacaccaaaaaaaatagcat-tggagaaaaatttt 948
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     91 AGTCTGAAGTTTACCTCATCAGTCCTGACACCAAAAAAATAGCATNTGGAGAAAAATTTT 150

Qy    949 aaggagatatccttttgctctcagggcatcagacacgtggaccactttgggtttatctgt 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151 AAGGAGATATCCTTTTGCTCTCAGGGCATCAGACACGTGGACCACTTTGGGTTTATCTGT 210

Qy   1009 cgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaat 1068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211 CGGGAGTCTTCCGGAGGTGGCGGCTTTCATTTTGTCTGTTACGTGTTTCAGTGCACAAAT 270

Qy   1069 gaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtg 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271 GAGGCTCTGGTTGATGAAATTATGATGACCCTGAAACAGGCCTTCACGGTGGCCGCAGTG 330

Qy   1129 cagcagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcac 1188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    331 CAGCAGACAGCTAAGGCGCCAGCCCAGCTGTGTGAGGGCTGCCCCCTGCAAAGCCTGCAC 390

Qy   1189 aagctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaag 1248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 AAGCTCTGTGAGAGGATAGAGGGAATGAATTCTTCCAAAACAAAACTAGAACTGCAAAAG 450

Qy   1249 cacctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattg 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 CACCTGACGACATTAACCAATCAGGAGCAGGCGACTATTTTTGAAGAGGTTCAGAAATTG 510

Qy   1309 agaccgagaaatgagcagcgagagaatgaattgattatttctttttctgagatgtttatat 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 AGACCGAGAAATGAGCAGCGAGAGAATGAATTGATTATTTCTTTTCTGAGATGTTTATAT 570

Qy   1369 gaagagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggca 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 GAAGAGAAACAGAAAGAACACATCCATATTGGGGAGATGAAGCAGACATCGCAGATGGCA 630

Qy   1429 gcagagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctg 1488
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    631 GCAGAGAATATTGGAAGTGAATTACCACCCAGTGCCCCTCGATTTAGGCTAGATATGCTG 690

Qy   1489 aaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaat 1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    691 AAAAACAAAGCAAAGAGATCTTTAACAGAGTCTTTAGAAAGTATTTTGTCCCGGGGTAAT 750

Qy   1549 aaagccagaggcctgcaggaacactccatcagtgt---ggatctggatagctcccct-gtc 1604
          |||||||||||||||||| |  | |      ||   |||||||||||||||||| |||
Db    751 AAAGCCAGAGGCCTGCAGGGAACCCTCCATCCCGTGTGGGATCTGGATAGCTCCCTGGTC 810

Qy   1605 tagtaca-ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggagg 1654
          ||||||| ||||||||||||||||||     |      |||||| |||   ||
Db    811 TAGTACATTTAAGTAACACCAGCAAAAGAGCCATCCTGTGTGTAAAAAGGG 861
```

```
RESULT   8
BM460573
LOCUS       BM460573                1109 bp    mRNA    linear   EST 05-FEB-2002
```

```
DEFINITION  AGENCOURT_6421241 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5531908
            5', mRNA sequence.
ACCESSION   BM460573
VERSION     BM460573.1  GI:18509613
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1109)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12214  row: f  column: 05
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .1109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5531908"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT      301 a    311 c    258 g    239 t
ORIGIN
```

```
  Query Match           20.5%;  Score 729.8;  DB 10;  Length 1109;
  Best Local Similarity   97.5%;  Pred. No. 2.2e-171;
  Matches  773;  Conservative    0;  Mismatches   17;  Indels    3;  Gaps    3;

Qy   1340 tgattatttcttttctgagatgtttatatgaagagaaacagaaagaacacatccatattg 1399
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     13 TGATTATTTCTTTTCTGAGATGTTTATATGAAGAGAAACAGAAAGAACACATCCATATTG 72

Qy   1400 gggagatgaagcagacatcgcagatggcagcagagaatattggaagtgaattaccaccca 1459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     73 GGGAGATGAAGCAGACATCGCAGATGGCAGCAGAGAATATTGGAAGTGAATTACCACCCA 132

Qy   1460 gtgccactcgatttaggctagatatgctgaaaaacaaagcaaagagatctttaacagagt 1519
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 GTGCCACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAGAGATCTTTAACAGAGT 192

Qy   1520 ctttagaaagtattttgtcccggggtaataaagccagaggcctgcaggaacactccatca 1579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 CTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTGCAGGAACACTCCATCA 252

Qy   1580 gtgtggatctggatagctccctgtctagtacattaagtaacaccagcaaagagccatctg 1639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 GTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCATCTG 312

Qy   1640 tgtgtgaaaaggaggccttgcccatctctgagagctcctttaagctcctcggctcctcgg 1699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 TGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCCTCGG 372

Qy   1700 aggacctgtccagtgactcggagagtcatctcccagaagagccagctccgctgtcgcccc 1759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 AGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCGCCCC 432

Qy   1760 agcaggccttcaggaggcgagcaaacaccctgagtcacttccccatcgaatgccaggaac 1819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    433 AGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAGGAAC 492

Qy   1820 ctccacaacctgcccgggggtccccgggggtttcgcaaaggaaacttatgaggtatcact 1879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    493 CTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTATCACT 552

Qy   1880 cagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatcttg 1939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    553 CAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTTG 612

Qy   1940 gtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatattcc 1999
          ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db    613 GTGATTCTGGTGGGACTCCTGTGAAGACCCCGGAGCATTCCTGGAGGCAGCAGATATTCC 672

Qy   2000 tccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattcag 2059
          ||||||||||||||||||||||||||| ||||||||||||||||||||| |||||||| |
Db    673 TCCGAGTAGCCACCCCGCAGAAGGCGGGCGATTCTTCCAGCAGATATGGAGATTATTCCG 732

Qy   2060 agct-gggagagcttcccccacgatc-tcctttagaaccagtttgt-gaagatgggccct 2116
          |||| |||||||||||||| | | | ||| |||||||||||||| ||||||||| ||
Db    733 AGCTGGGGAGAGCTTCCCCCCCCAACTTCCCTTAGAACCAGTTTGTGGAAGATGGGGCCC 792

Qy   2117 ttggccccccacc 2129
          || |  | || ||
Db    793 TTTGGGCGCCTCC 805
```

```
RESULT   9
BE299948
LOCUS       BE299948                952 bp    mRNA    linear   EST 20-JUL-2000
DEFINITION  600944494F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960782 5',
            mRNA sequence.
ACCESSION   BE299948
VERSION     BE299948.1  GI:9183696
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 952)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM54  row: k  column: 23
            High quality sequence stop: 709.
FEATURES             Location/Qualifiers
     source          1. .952
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2960782"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      249 a    240 c    250 g    212 t      1 others
ORIGIN
```

```
  Query Match           20.3%;  Score 722.2;  DB 10;  Length 952;
  Best Local Similarity  97.2%;  Pred. No. 1.7e-169;
  Matches  766;  Conservative    0;  Mismatches   19;  Indels    3;  Gaps    3;

Qy   1879 tcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatctt 1938
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTT 60

Qy   1939 ggtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatattc 1998
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATTC 120

Qy   1999 ctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattca 2058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTCA 180

Qy   2059 gagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgggcccttt 2118
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCCTTT 240

Qy   2119 ggcccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaag 2178
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAAG 300

Qy   2179 gctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaagcc 2238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGCC 360

Qy   2239 tctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccctgt 2298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCCTGT 420

Qy   2299 cttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaaatt 2358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAAATT 480

Qy   2359 aagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcaccga 2418
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    481 AAGTTTGACATGGAAAAAATGCACTCGGCTG-TGGGCAAGGTGTGCCACGTCATCACCGA 539

Qy   2419 ggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcccagcaaa 2478
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 GGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGCAAA 599

Qy   2479 cagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcagcat 2538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAGCAT 659

Qy   2539 gcgattctta-ttgaccttgggcgaacctttcctacacacccatacttctctgcccagct 2597
          |||||||||| ||||||| ||||||||||||||||||||||||||||||||||||| | |
Db    660 GCGATTCTTATTTGACCTGNGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAG-T 718

Qy   2598 tggagcaggacagctatcgctttacaacattttgaaggcctactcacttctagaccagga 2657
          |||||||||||||||||||||||||||||||   ||| ||||||   |      |||||
Db    719 TGGAGCAGGACAGCTATCGCTTTACAACATTTGTGAGGGCTACTCCACTTCTAGACAGGA 778

Qy   2658 agtgggat 2665
          || ||
Db    779 GTGGGTAT 786


RESULT  10
BI907513
LOCUS       BI907513                838 bp    mRNA    linear   EST 16-OCT-2001
DEFINITION  603065640F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214945 5',
            mRNA sequence.
ACCESSION   BI907513
VERSION     BI907513.1  GI:16170347
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11539  row: o  column: 10
            High quality sequence start: 4
            High quality sequence stop: 780.
FEATURES             Location/Qualifiers
     source          1. .838
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5214945"
                     /clone_lib="NIH_MGC_118"
                     /tissue_type="leukocyte"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated  adult donors.  Library is oligo-dT primed
                     and directionally cloned  (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb.  Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
BASE COUNT      237 a    199 c    191 g    211 t
ORIGIN


  Query Match           20.1%;  Score 713.8;  DB 10;  Length 838;
  Best Local Similarity  97.0%;  Pred. No. 2e-167;
  Matches  812;  Conservative    0;  Mismatches   17;  Indels    8;  Gaps    8;

Qy   2226 gaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaaga 2285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 GAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGA 68

Qy   2286 aattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactccagg 2345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 AATTACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGG 128

Qy   2346 aagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgcc 2405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 AAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCC 188

Qy   2406 acgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacacca 2465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 ACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCA 248

Qy   2466 gtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctgac 2525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 GTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGAC 308

Qy   2526 ttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatactt 2585
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    309 TTCCCAGCAGCATGCGATTCTTA-TGACCTTGGGCGAACCTTTCCTACACACCCATACTT 367

Qy   2586 ctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcact 2645
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    368 CTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACA-TTTGAAGGCCTACTCACT 426

Qy   2646 tctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttct 2705
          |||||||||||||||||||| |||||||||||||||| |||||||||| |||||||||
Db    427 TCTAGACCAGGAAGTGGGATA-TGCCAAGGTCTCAGC-TTGTAGCAGGCA-TTTGCTTCT 483
```

```
Qy   2706 tcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggct 2765
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    484 TCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATG-TTGACATGGGGCT 542

Qy   2766 gcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcgag 2825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 GCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGAG 602

Qy   2826 gttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggccc 2885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 GTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCC 662

Qy   2886 cagcctctacgctg-ccccctggttcctcaccatgtttgcctcacagttcccgctgggat 2944
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    663 CAGCCTCTACGCTGCCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGAT 722

Qy   2945 tcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtgg 3004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    723 TCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGG 782

Qy   3005 ctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaa 3061
          |||||  ||   ||||||| |||||||||  |        ||||||||||||||
Db    783 -TTTAAATCCTGTGGGAAGGCATAAGCCCTGATTCTGCAGCATTGAAAACCTAGAAA 838


RESULT  11
BG120498
LOCUS       BG120498                807 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602346815F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4441923 5',
            mRNA sequence.
ACCESSION   BG120498
VERSION     BG120498.1  GI:12614007
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10213  row: n  column: 04
            High quality sequence stop: 682.
FEATURES             Location/Qualifiers
     source          1. .807
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4441923"
                     /clone_lib="NIH_MGC_90"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      231 a    197 c    189 g    190 t
ORIGIN


  Query Match           19.8%;  Score 705;  DB 10;  Length 807;
  Best Local Similarity   96.1%;  Pred. No. 3.1e-165;
  Matches  766;  Conservative    0;  Mismatches   25;  Indels    6;  Gaps    4;

Qy   1992 gatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaaga 2051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGA 60

Qy   2052 ttattcagagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgg 2111
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGG 120

Qy   2112 gccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtg 2171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCTTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTG 180

Qy   2172 gcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagct 2231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCT 240

Qy   2232 ccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattac 2291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTAC 300

Qy   2292 tccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatc 2351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATC 360

Qy   2352 aaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtca 2411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCA 420

Qy   2412 tcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcc 2471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCC 480

Qy   2472 cagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttccca 2531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCA 540

Qy   2532 gcagcatgcgattcttattgaccttgggcgaacctttcctacaca-cccatacttctctg 2590
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    541 GCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACAGCCCATACTTCTCTG 600

Qy   2591 cccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcacttctag 2650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAG 660

Qy   2651 accaggaagtgggatattgcc--aaggtctcagcttt-gtagcaggcattttgcttcttc 2707
          |||||||||||||||||||  |||   |||||||| || |||||| | |||||||  ||
Db    661 ACCAGGAAGTGGGATATTGCCCAAAGTCTTCAGCTTTGGTGGCAGGCCTCTTGCTTCCTC 720

Qy   2708 atatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggctgc 2767
          |||||||||| |||||||||||  |||     ||   ||| ||||||||||||||| |
Db    721 ATATGAGTGACGAAGAGGCGTTAAAATGCTCCA--GTTCCGATGTTTGACATGGGCCGCG 778

Qy   2768 ggaaacagtatcggcca 2784
          |||| ||||   ||| |
Db    779 GGAACCAGTTCGGGCAA 795


RESULT  12
BM014940
LOCUS       BM014940                832 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603640877F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5416818 5',
            mRNA sequence.
ACCESSION   BM014940
VERSION     BM014940.1  GI:16529294
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 832)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12063  row: j  column: 19
            High quality sequence stop: 811.
FEATURES            Location/Qualifiers
     source         1. .832
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:5416818"
                    /clone_lib="NIH_MGC_87"
                    /tissue_type="mammary adenocarcinoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                    Average insert size 1.383 kb. Library enriched for
                    full-length clones and constructed by Life Technologies.
                    Note: this is a NIH_MGC Library."
BASE COUNT      231 a    205 c    186 g    208 t      2 others
ORIGIN


  Query Match             19.7%;  Score 701;  DB 10;  Length 832;
  Best Local Similarity   97.5%;  Pred. No. 3.1e-164;
  Matches  785;  Conservative    0;  Mismatches   12;  Indels    8;  Gaps    7;

Qy    2405 cacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacacc 2464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29 CACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACC 88

Qy    2465 agtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctga 2524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89 AGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGA 148

Qy    2525 cttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatact 2584
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     149 CTTCCCAGCAGCATGCGATTCTTA-TGACCTTGGGCGAACCTTTCCTACACACCCATACT 207

Qy    2585 tctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactcac 2644
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 TCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCAC 267

Qy    2645 ttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttc 2704
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     268 TTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTNTGCTTC 327

Qy    2705 ttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggc 2764
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 TTCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGC 387

Qy    2765 tgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcga 2824
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 TGCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGA 447

Qy    2825 ggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggcc 2884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 GGTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCC 507

Qy    2885 ccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctgggat 2944
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 CCAGCCTCTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGAT 567

Qy    2945 tcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaa-agtg 3003
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||| |||
Db     568 TCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAACAGTG 627

Qy    3004 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 3063
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 GCTTTAAGTCTGNTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACC 687

Qy    3064 atagttgacttttat-aaaaagcacgctacccaaccttggc-ttggtacagat-ggaaaag 3120
           |||||||||||||| |||||||||||||| |||| |||||| |||||||||| |||||||
Db     688 ATAGTTGACTTTATAAAAAAGCACGCTAACCAAGCTTGGCTTTGGTACAGATGGGAAAAG 747

Qy    3121 accatcaatcaggtatttgaaatggacatcgctaaa--cagttacaagcttatgaagttg 3178
           |||||||||||||||||| |||||||||||| ||||     |   ||||||||||||||||
Db     748 ACCATCAATCAGGTATTGGAAATGGACATCGGTAAAACAGTTACAAAGCTTATGAAGTTG 807

Qy    3179 agtacc-acgtccttcaagaagaac 3202
           |||||| ||||||||||||||||||
Db     808 AGTACCAACGTCCTTCAAGAAGAAC 832


RESULT  13
AL570425
LOCUS       AL570425                851 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL570425 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI020YF02 5
            prime, mRNA sequence.
ACCESSION   AL570425
VERSION     AL570425.1  GI:12926720
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 851)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES            Location/Qualifiers
     source         1. .851
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="CS0DI020YF02"
                    /clone_lib="LTI_NFL006_PL2"
                    /tissue_type="placenta"
                    /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                    was primed with a NotI-oligo(dT) primer.  Five prime end
                    enriched, double-stranded cDNA was digested with Not I and
                    cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                    vector. Library was normalized. Library was constructed by
                    Life Technologies. Contact : Feng Liang Life Technologies,
                    a division of Invitrogen 9800 Medical Center Drive
                    Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                    Email : fliang@lifetech.com URL :
                    http://fulllength.invitrogen.com"
BASE COUNT      240 a    214 c    220 g    174 t      3 others
ORIGIN


  Query Match             19.4%;  Score 689.8;  DB 9;  Length 851;
  Best Local Similarity   98.9%;  Pred. No. 2e-161;
  Matches  714;  Conservative    1;  Mismatches    4;  Indels    3;  Gaps    2;

Qy    2836 gattaccacagagacctctacaat--cacctggaggagcacgagatcggccccagcctct 2893
           |||||||||||||||||||||||     | ||||||||||||||||||||||||||||||
Db       1 GATTACCACAGAGACCTCTACAANTCACCTTGGAGGAGCACGAGATCGGCCCCAGCCTCT 60

Qy    2894 acgctgccccctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagcca 2953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCA 120

Qy    2954 gagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtc 3013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTC 180
```

```
Qy   3014 tgttgggaagccataagcccttgattctgcagcatgaaaacctagaaaccatagttgact 3073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACT 240

Qy   3074 ttataaaaagcacgctacccaaccttggcttggtacagatggaaaagaccatcaatcagg 3133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGG 300

Qy   3134 tatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttc 3193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TATTTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTC 360

Qy   3194 aagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagaga 3253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGA 420

Qy   3254 aaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaa 3313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAACCAACAGCAGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAA 480

Qy   3314 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 3373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGGTAGGATCCAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGC 540

Qy   3374 tgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggagg 3433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TGAAGCAGGCCATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGG 600

Qy   3434 agctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagc 3493
          |||||||||||||| :||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGCTGCGGCGGCGGA-SGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGC 659

Qy   3494 ccacgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcc 3553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 CCACGGGCGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCC 719

Qy   3554 tt 3555
          ||
Db    720 TT 721


RESULT  14
BE300005
LOCUS       BE300005               754 bp    mRNA    linear   EST 20-JUL-2000
DEFINITION  600944594F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960783 5',
            mRNA sequence.
ACCESSION   BE300005
VERSION     BE300005.1  GI:9183753
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM54  row: k  column: 24
            High quality sequence start: 2
            High quality sequence stop: 753.
FEATURES             Location/Qualifiers
     source          1. .754
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2960783"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      227 a    186 c    177 g    164 t
ORIGIN


  Query Match             19.2%;  Score 683.4;  DB 10;  Length 754;
  Best Local Similarity   99.1%;  Pred. No. 7.5e-160;
  Matches  750;  Conservative    0;  Mismatches    1;  Indels    6;  Gaps    6;

Qy   1878 ctcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatct 1937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATC- 59

Qy   1938 tggtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatatt 1997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 TGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATT 119

Qy   1998 cctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattc 2057
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTC 179

Qy   2058 agagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgggccctt 2117
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||| |
Db    180 AGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAG-TTGTGAAGATGGGCCC-T 237

Qy   2118 tggccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaa 2177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    238 TGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAA 297

Qy   2178 ggctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaagc 2237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    298 GGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGC 357

Qy   2238 ctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccctg 2297
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    358 CTCTGAAAATGA-TTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCCTG 416

Qy   2298 tcttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaaat 2357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 TCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAAAT 476

Qy   2358 taagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcaccg 2417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 TAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCATCACCG 536

Qy   2418 aggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcccagcaa 2477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537 AGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGCAA 596

Qy   2478 acagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcagca 2537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    597 ACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAGCA 656

Qy   2538 tgcgattcttattgaccttgggcgaacctttcctacacacccatacttctctgcccagct 2597
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    657 TGCGATTCTTA-TGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAGCT 715

Qy   2598 tggagcaggacagctatcgctttacaacattttgaag 2634
          |||||||||||||||||| |||||||||||||| |
Db    716 TGGAGCAGGACAGCTATCGC-TTACAACATTTTGAGG 751


RESULT  15
```

```
BE561889
LOCUS       BE561889                892 bp    mRNA    linear   EST 15-AUG-2000
DEFINITION  601346094F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3678887 5',
            mRNA sequence.
ACCESSION   BE561889
VERSION     BE561889.1  GI:9805609
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM356  row: l  column: 24
            High quality sequence start: 50
            High quality sequence stop: 758.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3678887"
                     /clone_lib="NIH_MGC_8"
                     /tissue_type="Burkitt lymphoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      224 a    240 c    242 g    186 t
ORIGIN
```

Query Match 19.2%; Score 682.6; DB 10; Length 892;
Best Local Similarity 95.3%; Pred. No. 1.3e-159;
Matches 790; Conservative 0; Mismatches 29; Indels 10; Gaps 8;

```
Qy   1628 aagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaagctcc 1687
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     63 AAGAGCCATCTGTGTGTGAAAAGGAGG-CTTGCCCATCTCTGAGAGCTCCTTTAAGCTCC 121

Qy   1688 tcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagccagctc 1747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 TCGGCTCCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTC 181

Qy   1748 cgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcactt-ccccatc 1806
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    182 CGCTGTCGCCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCCATC 241

Qy   1807 gaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaacttt 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 GAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTT 301

Qy   1867 atgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaa 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 ATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAA 361

Qy   1927 gcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctggagg 1986
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 GCAAATTATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGG 421

Qy   1987 cagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatat 2046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 CAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATAT 481

Qy   2047 gaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgtgaa 2106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 GAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAA 541

Qy   2107 gatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccgagag 2166
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    542 GATGGGCCCTTTGGCCCCCCACCAGAGGAAACGAAAAGGACATCTCGTGAGCTCCGAGAG 601

Qy   2167 ctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaatcag 2226
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    602 CTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGCGAAGGAAAATCAG 661

Qy   2227 aagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaa 2286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 AAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAA 721

Qy   2287 attactccctg-tcttaaagaagtaactacagtgtggga--aaagatgcttagcactcca 2343
          |||||||||||| ||||||||||||||||||||||||||  ||    ||||||||||||
Db    722 ATTACTCCCTGTTCTTAAAGAAGTAACTACAGTGTGGGACCAAGATGGCTTAGCACTCCA 781

Qy   2344 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 2403
          ||||||||    ||| |||||| | |||  | |||||  |||||  |||||||| ||| | |
Db    782 GGAAGATC-CCCATTTAGTTTGGCCTGGTGACAATGC-CTCGGGTGTTGGGCCAGGGGGG 839

Qy   2404 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattcc 2452
          || |||||||||| | |||||   ||  | || |||||| ||||||||
Db    840 CC-CGTCATCACCGCGCTGAAACTGGG--AGTTTTAGCTGGGCAATTCC 885
```

Search completed: September 9, 2002, 16:41:06
Job time: 9234 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 14:27:02 ; Search time 706.85 Seconds
(without alignments)
8634.973 Million cell updates/sec

Title: US-09-762-311-3_COPY_171_3725
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3555	100.0	3983	21	AAZ50906	Human TBC-1 cDNA f
2	3555	100.0	3988	21	AAZ50907	Human TBC-1 cDNA f
3	3026.4	85.1	3326	22	AAI57854	Human polynucleoti
4	2940.2	82.7	3727	21	AAC77404	Human ORFX ORF2959
5	2833.8	79.7	3281	22	AAI59640	Human polynucleoti
6	2510.4	70.6	4039	19	AAV05886	Human Tbc-1 gene.
7	2112	59.4	2362	22	AAH18401	Human cDNA sequenc
8	1000.8	28.2	1422	22	AAS34787	cDNA encoding nove
9	999.4	28.1	2989	22	AAK94796	Human full-length
10	943.2	26.5	1269	22	AAS34785	cDNA encoding nove
11	785	22.1	883	22	AAH04450	Human cDNA clone (
12	574.2	16.2	849	22	AAK92006	Human cDNA 5'-end
13	574.2	16.2	849	22	AAK93666	Human cDNA clone r
14	422.6	11.9	17590	21	AAZ50904	Human TBC-1 partia
15	362	10.2	696	20	AAZ15229	Human gene express
16	306.2	8.6	5472	23	ABL02013	Drosophila melanog
17	284.4	8.0	99960	21	AAZ50905	Human TBC-1 partia
18	280	7.9	458	20	AAX40323	Human secreted pro
c 19	245.6	6.9	8105	23	ABL02012	Drosophila melanog
c 20	228.4	6.4	555	22	AAH09958	Human cDNA clone (
21	132.8	3.7	3691	22	AAF58616	Human RECAP polynu
22	132.8	3.7	4827	22	AAI59592	Human polynucleoti
23	132.8	3.7	5294	22	AAI57806	Human polynucleoti
24	128.8	3.6	3585	23	ABL12337	Drosophila melanog
25	123.8	3.5	154	23	AAS58636	cDNA #1312 encodin
26	119.4	3.4	2626	23	ABL18133	Drosophila melanog
27	119.4	3.4	7829	23	ABL18132	Drosophila melanog
28	108	3.0	2424	23	AAS71793	DNA encoding novel
29	83.8	2.4	1601	21	AAC33401	Arabidopsis thalia
c 30	79.6	2.2	6765	23	ABL12336	Drosophila melanog
31	78.4	2.2	300	20	AAZ13067	Human gene express
32	75.2	2.1	1492	21	AAC38750	Arabidopsis thalia
c 33	73.8	2.1	1140	21	AAZ97020	Human secreted pro
34	68.4	1.9	3342	22	AAF32789	Human secreted pro
35	68	1.9	2492	21	AAC78199	Human cancer assoc
36	65.2	1.8	1831	22	AAH64964	Human secreted pro
37	65	1.8	1117	22	AAI61302	Human polynucleoti
38	65	1.8	2035	22	AAF59600	Human cell cycle a
39	65	1.8	2482	22	AAI59516	Human polynucleoti
40	65	1.8	2666	22	AAH17866	Human cDNA sequenc
41	62.6	1.8	2531	23	ABL07359	Drosophila melanog
42	62.6	1.8	4661	23	ABL07358	Drosophila melanog
43	61.6	1.7	2801	22	AAF32643	Human cDNA encodin
c 44	61.6	1.7	2803	22	AAS27080	cDNA encoding nove
45	61.6	1.7	2903	22	AAK51641	Human polynucleoti

ALIGNMENTS

RESULT 1
AAZ50906
ID AAZ50906 standard; cDNA; 3983 BP.
XX
AC AAZ50906;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human TBC-1 cDNA from first transcript.
XX
KW TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;
KW Single nucleotide polymorphism; tissue differentiation; prostate cancer;
KW linkage analysis; genetic map; detection; diagnosis; genotyping;
KW transgenic animal; screening; alternative splicing; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..170
FT /*tag= a
FT /note= "Spans through exon 1 and part of exon 2"
FT CDS 171..3677
FT /*tag= b
FT /product= "Human TBC-1 protein"
FT /note= "First transcription product"
FT 3'UTR 3726..3983
FT /*tag= c
FT polyA_signal 3942..3947
FT /*tag= d
XX
PN WO200008209-A2.
XX

PD 17-FEB-2000.
XX
PF 06-AUG-1999; 99WO-IB01444.
XX
PR 07-AUG-1998; 98US-0095653.
XX
PA (GEST) GENSET.
XX
PI Blumenfeld M, Bougueleret L, Chumakov I;
XX
DR WPI; 2000-205736/18.
DR P-PSDB; AAY45096.
XX
PT New isolated human TBC-1 nucleic acids, useful for developing products
PT for the diagnosis and treatment of disorders involving cell
PT proliferation, particularly prostate cancer -
XX
PS Claim 3; Page 150-155; 166pp; English.
XX
CC The present sequence is the cDNA encoding the first transcription
CC product of human TBC-1 protein, comprising exons 1-2 and A-L.
CC Alternative splicing events result in two mRNA molecules, due to
CC splicing at two distinct first exons, exon1 and exon 1bis. TBC-1 gene
CC is mapped to a candidate region of prostate cancer on chromosome 4.
CC TBC-1 gene is involved in the regulation of cell cycle and tissue
CC differentiation in mammals. An alteration of TBC-1 sequence may be
CC associated with a pathological condition, resulting in abnormal cell
CC proliferation leading to cancer, e.g. prostate cancer. The biallelic
CC markers can be used for generation of genetic maps, linkage analysis and
CC association studies. TBC-1 sequence can be used for detection,
CC diagnosis, genotyping, production of transgenic animals and screening
CC of compounds for use in therapy.
XX
SQ Sequence 3983 BP; 1090 A; 1007 C; 1007 G; 878 T; 1 other;

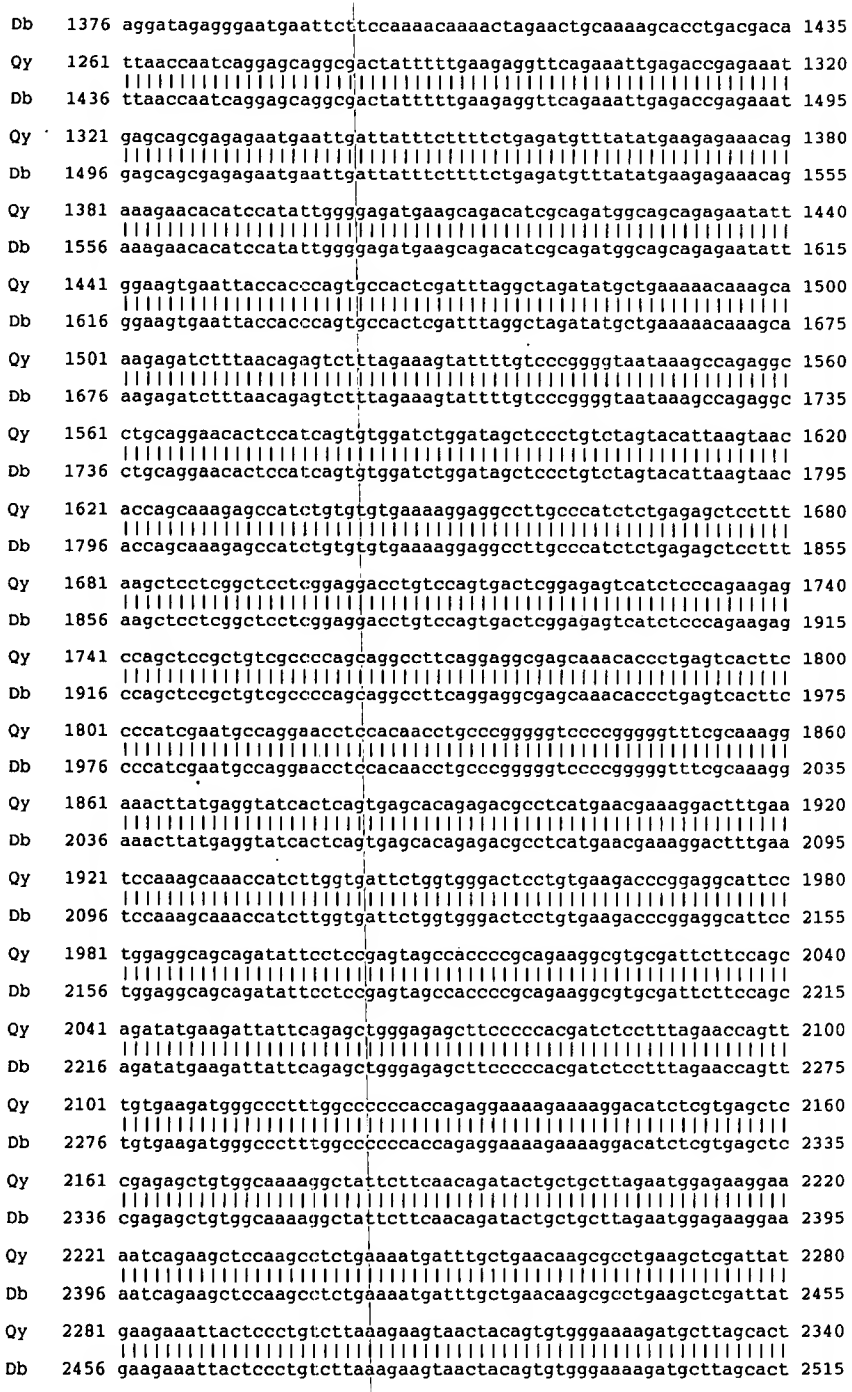
Query Match 100.0%; Score 3555; DB 21; Length 3983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 230

Qy     61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    231 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 290

Qy    121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    291 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 350

Qy    181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    351 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 410

Qy    241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    411 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 470

Qy    301 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    471 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 530

Qy    361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    531 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 590

Qy    421 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    591 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 650

Qy    481 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    651 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 710

Qy    541 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    711 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 770

Qy    601 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    771 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 830

Qy    661 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    831 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 890

Qy    721 ctgcgctcgctggcctttaggaaggagctgcaggatggggcctccgaagcagcggcttc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    891 ctgcgctcgctggcctttaggaaggagctgcaggatggggcctccgaagcagcggcttc 950

Qy    781 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    951 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 1010

Qy    841 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 1070

Qy    901 tacctcatcagtcctgacaccaaaaaaaatagcattggagaaaaatttttaaggagatatcc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 tacctcatcagtcctgacaccaaaaaaaatagcattggagaaaaatttttaaggagatatcc 1130

Qy    961 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1190

Qy   1021 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1250

Qy   1081 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1310

Qy   1141 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1370

Qy   1201 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1430

Qy   1261 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1490

Qy   1321 gagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaacag 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 gagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaagagaaacag 1550

Qy   1381 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1610

Qy   1441 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1670

Qy   1501 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1730

Qy   1561 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1731 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1790

Qy   1621 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1850

Qy   1681 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1910

Qy   1741 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1970

Qy   1801 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 2030

Qy   1861 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 2090

Qy   1921 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2091 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 2150

Qy   1981 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2151 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2210

Qy   2041 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2211 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2270

Qy   2101 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2271 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2330

Qy   2161 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2331 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2390

Qy   2221 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2391 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2450

Qy   2281 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2451 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2510

Qy   2341 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2511 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2570

Qy   2401 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2571 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2630

Qy   2461 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2631 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2690

Qy   2521 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2691 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2750

Qy   2581 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2751 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2810

Qy   2641 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2811 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2870

Qy   2701 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2871 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2930

Qy   2761 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2990

Qy   2821 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 3050

Qy   2881 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 3110

Qy   2941 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3170

Qy   3001 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3171 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3230

Qy   3061 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3290

Qy   3121 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3350

Qy   3181 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3410

Qy   3241 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3411 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3470

Qy   3301 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3471 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3530

Qy   3361 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3531 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3590

Qy   3421 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3591 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3650

Qy   3481 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3651 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3710

Qy   3541 cactgtccaggcctt 3555
          |||||||||||||||
Db   3711 cactgtccaggcctt 3725
```

RESULT 2
AAZ50907
ID AAZ50907 standard; cDNA; 3988 BP.
XX
AC AAZ50907;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human TBC-1 cDNA from second transcript.
XX
KW TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;

```
KW   Single nucleotide polymorphism; tissue differentiation; prostate cancer;
KW   linkage analysis; genetic map; detection; diagnosis; genotyping;
KW   transgenic animal; screening; alternative splicing; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..175
FT                   /*tag= a
FT   misc_feature    1..458
FT                   /*tag= b
FT                   /note= "Homologous to 5' EST obtained from human
FT                   pancreas cDNA library"
FT   CDS             176..3682
FT                   /*tag= c
FT                   /product= "Human TBC-1 protein"
FT                   /note= "Second transcription product"
FT   3'UTR           3731..3988
FT                   /*tag= d
FT   polyA_signal    3947..3952
FT                   /*tag= e
XX
PN   WO200008209-A2.
XX
PD   17-FEB-2000.
XX
PF   06-AUG-1999;   99WO-IB01444.
XX
PR   07-AUG-1998;   98US-0095653.
XX
PA   (GEST ) GENSET.
XX
PI   Blumenfeld M,  Bougueleret L,  Chumakov I;
XX
DR   WPI; 2000-205736/18.
DR   P-PSDB; AAY45096.
XX
PT   New isolated human TBC-1 nucleic acids, useful for developing products
PT   for the diagnosis and treatment of disorders involving cell
PT   proliferation, particularly prostate cancer -
XX
PS   Claim 4; Page 156-161; 166pp; English.
XX
CC   The present sequence is the cDNA encoding the second transcription
CC   product of human TBC-1 protein, comprising exons 1bis to 2 and A-L.
CC   Alternative splicing events result in two mRNA molecules, due to
CC   splicing at two distinct first exons, exon1 and exon 1bis. TBC-1 gene
CC   is mapped to a candidate region of prostate cancer on chromosome 4.
CC   TBC-1 gene is involved in the regulation of cell cycle and tissue
CC   differentiation in mammals. An alteration of TBC-1 sequence may be
CC   associated with a pathological condition, resulting in abnormal cell
CC   proliferation leading to cancer, e.g. prostate cancer. The biallelic
CC   markers can be used for generation of genetic maps, linkage analysis and
CC   association studies. TBC-1 sequence can be used for detection,
CC   diagnosis, genotyping, production of transgenic animals and screening
CC   of compounds for use in therapy.
XX
SQ   Sequence 3988 BP; 1107 A; 983 C; 1002 G; 896 T; 0 other;


  Query Match             100.0%;  Score 3555;  DB 21;  Length 3988;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3555;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 235

Qy     61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    236 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 295

Qy    121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    296 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 355

Qy    181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    356 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 415

Qy    241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    416 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 475

Qy    301 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    476 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 535

Qy    361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    536 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 595

Qy    421 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    596 cctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctgcac 655

Qy    481 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    656 tgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggccgc 715

Qy    541 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    716 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 775

Qy    601 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    776 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 835

Qy    661 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    836 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 895

Qy    721 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    896 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 955

Qy    781 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    956 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 1015

Qy    841 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1016 cagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaagtt 1075

Qy    901 tacctcatcagtcctgacaccaaaaaaaatagcattggagaaaaatttttaaggagatatcc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1076 tacctcatcagtcctgacaccaaaaaaaatagcattggagaaaaatttttaaggagatatcc 1135

Qy    961 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1136 ttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtcttcc 1195

Qy   1021 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1196 ggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgaggctctggtt 1255

Qy   1081 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1256 gatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcagcagacagct 1315

Qy   1141 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1316 aaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaagctctgtgag 1375

Qy   1201 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 1376 aggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcacctgacgaca 1435
Qy 1261 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1320
Db 1436 ttaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgagaccgagaaat 1495
Qy 1321 gagcagcgagagaatgaattgattatttctttttctgagatgtttatatgaagagaaacag 1380
Db 1496 gagcagcgagagaatgaattgattatttctttttctgagatgtttatatgaagagaaacag 1555
Qy 1381 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1440
Db 1556 aaagaacacatccatattggggagatgaagcagacatcgcagatggcagcagagaatatt 1615
Qy 1441 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1500
Db 1616 ggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagca 1675
Qy 1501 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1560
Db 1676 aagagatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggc 1735
Qy 1561 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1620
Db 1736 ctgcaggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaac 1795
Qy 1621 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1680
Db 1796 accagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttt 1855
Qy 1681 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1740
Db 1856 aagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagag 1915
Qy 1741 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1800
Db 1916 ccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttc 1975
Qy 1801 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 1860
Db 1976 cccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaagg 2035
Qy 1861 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 1920
Db 2036 aaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaa 2095
Qy 1921 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 1980
Db 2096 tccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcc 2155
Qy 1981 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2040
Db 2156 tggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagc 2215
Qy 2041 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2100
Db 2216 agatatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtt 2275
Qy 2101 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2160
Db 2276 tgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatctcgtgagctc 2335
Qy 2161 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2220
Db 2336 cgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaa 2395
Qy 2221 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2280
Db 2396 aatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattat 2455
Qy 2281 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2340
Db 2456 gaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcact 2515
Qy 2341 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2400
Db 2516 ccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggt 2575
Qy 2401 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2460
Db 2576 gtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaa 2635
Qy 2461 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2520
Db 2636 caccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcag 2695
Qy 2521 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2580
Db 2696 ctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacaccca 2755
Qy 2581 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2640
Db 2756 tacttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctac 2815
Qy 2641 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2700
Db 2816 tcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttg 2875
Qy 2701 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2760
Db 2876 cttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatg 2935
Qy 2761 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2820
Db 2936 gggctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctc 2995
Qy 2821 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 2880
Db 2996 tcgaggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatc 3055
Qy 2881 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 2940
Db 3056 ggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctg 3115
Qy 2941 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3000
Db 3116 ggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaa 3175
Qy 3001 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3060
Db 3176 gtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaa 3235
Qy 3061 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3120
Db 3236 accatagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaag 3295
Qy 3121 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3180
Db 3296 accatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgag 3355
Qy 3181 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3240
Db 3356 taccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatg 3415
Qy 3241 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3300
Db 3416 gataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacag 3475
Qy 3301 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3360
Db 3476 ttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagc 3535
Qy 3361 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3420
Db 3536 agtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctg 3595

```
Qy   3421 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3596 cagacggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgc 3655

Qy   3481 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3656 acgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatccca 3715

Qy   3541 cactgtccaggcctt 3555
          |||||||||||||||
Db   3716 cactgtccaggcctt 3730


RESULT   3
AAI57854
ID   AAI57854 standard; cDNA; 3326 BP.
XX
AC   AAI57854;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 57.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM38698.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -  .
XX
PS   Claim 1; SEQ ID NO 57; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3326 BP; 924 A; 824 C; 843 G; 735 T; 0 other;


  Query Match             85.1%;  Score 3026.4;  DB 22;  Length 3326;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3030;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    520 gaggtgctcttctgcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatc 579
          | | | ||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db     43 gggatcctctagagcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatc 102

Qy    580 gacgagtgcatcgagaagttcaatcacgtcagcggcagccgggggtccgagagccccccgc 639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103 gacgagtgcatcgagaagttcaatcacgtcagcggcagccgggggtccgagagccccccgc 162

Qy    640 cccaacccgccccatgccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgc 699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    163 cccaacccgccccatgccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgc 222

Qy    700 aagtccttctcccagcccggcctgcgctcgctggcctttaggaaggagctgcaggatggg 759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    223 aagtccttctcccagcccggcctgcgctcgctggcctttaggaaggagctgcaggatggg 282

Qy    760 ggcctccgaagcagcggcttcttcagctccttcgaggagagcgacattgagaaccacctc 819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    283 ggcctccgaagcagcggcttcttcagctccttcgaggagagcgacattgagaaccacctc 342

Qy    820 attagcggacacaatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttc 879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    343 attagcggacacaatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttc 402

Qy    880 acgattggccagtctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggag 939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    403 acgattggccagtctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggag 462

Qy    940 aaaaattttaaggagatatccttttgctctcagggcatcagacacgtggaccactttggg 999
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    463 aaaaattttaaggagatatccttttgctctcagggcatcagacacgtggaccactttggg 522

Qy   1000 tttatctgtcgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcag 1059
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    523 tttatctgtcgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcag 582

Qy   1060 tgcacaaatgaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtg 1119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    583 tgcacaaatgaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtg 642

Qy   1120 gccgcagtgcagcagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaa 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    643 gccgcagtgcagcagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaa 702

Qy   1180 agcctgcacaagctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaa 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    703 agcctgcacaagctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaa 762

Qy   1240 ctgcaaaagcacctgacgacattaaccaatcaggagcaggcgactatttttgaagaggtt 1299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    763 ctgcaaaagcacctgacgacattaaccaatcaggagcaggcgactatttttgaagaggtt 822

Qy   1300 cagaaattgagaccgagaaatgagcagcgagagaatgaattgattatttcttttctgaga 1359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    823 cagaaattgagaccgagaaatgagcagcgagagaatgaattgattatttcttttctgaga 882

Qy   1360 tgtttatatgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcg 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    883 tgtttatatgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcg 942

Qy   1420 cagatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggcta 1479
```

```
Db    943 cagatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggcta 1002

Qy   1480 gatatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcc 1539

Db   1003 gatatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcc 1062

Qy   1540 cggggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctcc 1599

Db   1063 cggggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctcc 1122

Qy   1600 ctgtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttg 1659

Db   1123 ctgtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttg 1182

Qy   1660 cccatctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcg 1719

Db   1183 cccatctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcg 1242

Qy   1720 gagagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcga 1779

Db   1243 gagagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcga 1302

Qy   1780 gcaaacaccctgagtcacttccccatcgaatgccaggaacctccacaacctgcccggggg 1839

Db   1303 gcaaacaccctgagtcacttccccatcgaatgccaggaacctccacaacctgcccggggg 1362

Qy   1840 tccccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcct 1899

Db   1363 tccccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcct 1422

Qy   1900 catgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcct 1959

Db   1423 catgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcct 1482

Qy   1960 gtgaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcag 2019

Db   1483 gtgaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcag 1542

Qy   2020 aaggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttccccca 2079

Db   1543 aaggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttccccca 1602

Qy   2080 cgatctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaag 2139

Db   1603 cgatctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaag 1662

Qy   2140 aaaaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg 2199

Db   1663 aaaaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg 1722

Qy   2200 ctgcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaac 2259

Db   1723 ctgcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaac 1782

Qy   2260 aagcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtg 2319

Db   1783 aagcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtg 1842

Qy   2320 tgggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatg 2379

Db   1843 tgggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatg 1902

Qy   2380 cactcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttcta 2439

Db   1903 cactcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttcta 1962

Qy   2440 gctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgcca 2499

Db   1963 gctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgcca 2022

Qy   2500 tacaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttggg 2559

Db   2023 tacaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttggg 2082

Qy   2560 cgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctt 2619

Db   2083 cgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctt 2142

Qy   2620 tacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctc 2679

Db   2143 tacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctc 2202

Qy   2680 agctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctc 2739

Db   2203 agctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctc 2262

Qy   2740 aagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattatttta 2799

Db   2263 aagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattatttta 2322

Qy   2800 cagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaat 2859

Db   2323 cagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaat 2382

Qy   2860 cacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatg 2919

Db   2383 cacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatg 2442

Qy   2920 tttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcag 2979

Db   2443 tttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcag 2502

Qy   2980 ggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgatt 3039

Db   2503 ggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgatt 2562

Qy   3040 ctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaacctt 3099

Db   2563 ctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaacctt 2622

Qy   3100 ggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacag 3159

Db   2623 ggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacag 2682

Qy   3160 ttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcct 3219

Db   2683 ttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcct 2742

Qy   3220 ctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacag 3279

Db   2743 ctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacag 2802

Qy   3280 aaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcc 3339

Db   2803 aaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcc 2862

Qy   3340 accattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaa 3399

Db   2863 accattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaa 2922

Qy   3400 ctggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagccc 3459

Db   2923 ctggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagccc 2982

Qy   3460 agcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagg 3519

Db   2983 agcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagg 3042

Qy   3520 agagattgcaacaccatcccacactgtccaggcctt 3555

Db   3043 agagattgcaacaccatcccacactgtccaggcctt 3078


RESULT   4
AAC77404
```

ID AAC77404 standard; cDNA; 3727 BP.
XX
AC AAC77404;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2959 polynucleotide sequence SEQ ID NO:5917.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW- immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB43195.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 5; Page 5094-5096; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 3727 BP; 1031 A; 889 C; 931 G; 875 T; 1 other;

Query Match 82.7%; Score 2940.2; DB 21; Length 3727;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 3005; Conservative 0; Mismatches 13; Indels 7; Gaps 5;

```
Qy   532 tgcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatc 591
Db     4 tgcggccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatc 63

Qy   592 gagaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccc 651
Db    64 gagaagttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccc 123

Qy   652 catgccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcc 711
Db   124 catgccgcgcgccacagggagccaggacctgtgcgcaggcccatgcgcaagtccttctcc 183

Qy   712 cagcccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagc 771
Db   184 cagcccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagc 243

Qy   772 agcggcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacac 831
Db   244 agcggcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacac 303

Qy   832 aatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccag 891
Db   304 aatattgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccag 363

Qy   892 tctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaattttaag 951
Db   364 tctgaagtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaattttaag 423

Qy   952 gagatatccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgg 1011
Db   424 gagatatccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgg 483

Qy  1012 gagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
Db   484 gagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 543

Qy  1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
Db   544 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 603

Qy  1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
Db   604 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 663

Qy  1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
Db   664 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 723

Qy  1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
Db   724 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 783

Qy  1312 ccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaa 1371
Db   784 ccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaa 843

Qy  1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
Db   844 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 903

Qy  1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
Db   904 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 963

Qy  1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1551
Db   964 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1023
```

```
Qy   1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1083

Qy   1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1084 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1143

Qy   1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1203

Qy   1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1204 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1263

Qy   1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1323

Qy   1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1383

Qy   1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1443

Qy   1972 aggcattcctggaggcagcagatattcctccgagtagccacccccgcagaaggcgtgcgat 2031
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 aggcattcctggaggcagcagatattcctccgagtagccacccccgcagaaggcgtgcgat 1503

Qy   2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 1563

Qy   2092 gaaccagtttgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggacatct 2151
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   1564 gaaccagtttgtgaagatgggccctttgg-cccccaccagaggaaaagaaaaggacatct 1622

Qy   2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgc-tgcttagaat 2210
          |||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||
Db   1623 cgtgagctccgagagctgtggcaaagggctattcttcaacagatactgcntgcttagaat 1682

Qy   2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1683 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 1742

Qy   2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1743 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 1802

Qy   2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1803 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 1862

Qy   2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1863 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 1922

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1923 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 1982

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1983 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2042

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2043 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2102

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2103 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2162

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2163 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2222

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2223 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2282

Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2283 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2342

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2343 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2402

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2403 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2462

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2463 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 2522

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2523 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 2582

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2583 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 2642

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2643 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 2702

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2703 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 2762

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2763 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 2822

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2823 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 2882

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
          |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2883 g--cctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 2940

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcaggagagattgcaa 3530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 gcctgagtgcacgcagcccgagcccacgggcgactgacagct-tgcaggagagattgcaa 2999

Qy   3531 caccatcccacactgtccaggcctt 3555
          ||||||  ||||||||||||||||
Db   3000 caccat--cacactgtccaggcctt 3022
```

```
RESULT   5
AAI59640
ID   AAI59640 standard; cDNA; 3281 BP.
XX
AC   AAI59640;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 3629.
```

```
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM40484.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 3629; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 3281 BP; 912 A; 812 C; 824 G; 733 T; 0 other;


  Query Match             79.7%;  Score 2833.8;  DB 22;  Length 3281;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2983;  Conservative    0;  Mismatches   32;  Indels   20;  Gaps   11;

Qy    541 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaagttc 60

Qy    601 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgccgcg 120

Qy    661 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    121 cccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagcccggc 180

Qy    721 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggcttc 240

Qy    781 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatattgtg 300

Qy    841 cagcccacagatat--cgaggaaaatcgaactat---gctcttcacgattggccagtctg 895
          ||||||||||||||     | ||| |||||||||    || |      | ||||||||||
Db    301 cagcccacagatattcgaggaaaattcgaactattgctctttaacgattgggccagtctg 360

Qy    896 aagtttacctcatcagtcctgacaccaaaaaaat-----agcattggagaaaaattttaa 950
          ||||||||||||||||||||||||||||||||     |      || |||  ||| |
Db    361 aagtttacctcatcagtcctgacaccaaaaaaattaggcattgggagaaaaattttaag 420

Qy    951 ggagatatccttttgctctcagggcatcagaca-cgtggaccactttggg-tttatctgt 1008
          |||||||||||||||||||||||||||||||| |||||||||||| || |||||||||
Db    421 ggagatatccttttgctctcagggcatcagacaccgtggaccacttttggttttatctgt 480

Qy   1009 cgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaat 1068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 cgggagtcttccggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaat 540

Qy   1069 gaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtg 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gaggctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtg 600

Qy   1129 cagcagacagctaaggcgccagcccagctgtgtgagggctg--ccccctgcaaagcc--t 1184
          ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db    601 cagcagacagctaaggcgccagcccagctgtgtgagggctgccccccctgcaaagccctg 660

Qy   1185 gcacaagctctgtgagaggatagagggaatgaatt-cttccaaaacaaaactagaactgc 1243
          |||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db    661 ccacaagctctgtgagaggatagagggaatgaattcctcccaaaacaaaactagaactgc 720

Qy   1244 aaaagca-cctgacgacattaa-ccaatcaggagcaggcgactatttttgaagaggttca 1301
          ||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    721 aaaagcaccctgacgacattaacccaatcaggagcaggcgactatttttgaagaggttca 780

Qy   1302 gaaattgagaccgagaaatgagcagcgagagaatgaattgattatttctttttctgagatg 1361
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    781 gaaattgagaccgagaaatgagcagcgagagaatgaattgattatttctttttgagatg 840

Qy   1362 tttatatgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcgca 1421
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 tttatttgaagagaaacagaaagaacacatccatattggggagatgaagcagacatcgca 900

Qy   1422 gatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggctaga 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 gatggcagcagagaatattggaagtgaattaccacccagtgccactcgatttaggctaga 960

Qy   1482 tatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccg 1541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 tatgctgaaaaacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccg 1020

Qy   1542 gggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctccct 1601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gggtaataaagccagaggcctgcaggaacactccatcagtgtggatctggatagctccct 1080

Qy   1602 gtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcc 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 gtctagtacattaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcc 1140

Qy   1662 catctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcgga 1721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 catctctgagagctcctttaagctcctcggctcctcggaggacctgtccagtgactcgga 1200

Qy   1722 gagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagc 1781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 gagtcatctcccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagc 1260
```

```
Qy   1782 aaacaccctgagtcacttcccatcgaatgccaggaacctccacaacctgcccgggggtc 1841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 aaacaccctgagtcacttcccatcgaatgccaggaacctccacaacctgcccgggggtc 1320

Qy   1842 cccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctca 1901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 cccgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctca 1380

Qy   1902 tgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgt 1961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 tgaacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgt 1440

Qy   1962 gaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaa 2021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 gaagacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaa 1500

Qy   2022 ggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttcccccacg 2081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttcccccacg 1560

Qy   2082 atctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaa 2141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 atctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaa 1620

Qy   2142 aaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgct 2201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 aaggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgct 1680

Qy   2202 gcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaa 2261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 gcttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaa 1740

Qy   2262 gcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtg 2321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 gcgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtg 1800

Qy   2322 ggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgca 2381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ggaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgca 1860

Qy   2382 ctcggctgttgggcaaggtgtg-ccacgtcatcaccgaggtgaaatctggaaatttctag 2440
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1861 ctcggctgttgggcaaggtgtgcccacgtcatcaccgaggtgaaatctggaaatttctag 1920

Qy   2441 ctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccat 2500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ctgagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccat 1980

Qy   2501 acaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggc 2560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 acaaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggc 2040

Qy   2561 gaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgcttt 2620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 gaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgcttt 2100

Qy   2621 acaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctca 2680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 acaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctca 2160

Qy   2681 gctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctca 2740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 gctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctca 2220

Qy   2741 agtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttac 2800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 agtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttac 2280

Qy   2801 agatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatc 2860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 agatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatc 2340

Qy   2861 acctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgt 2920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 acctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgt 2400

Qy   2921 ttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagg 2980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 ttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagg 2460

Qy   2981 gaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattc 3040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 gaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattc 2520

Qy   3041 tgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttg 3100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 tgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttg 2580

Qy   3101 gcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagt 3160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 gcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagt 2640

Qy   3161 tacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctc 3220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 tacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctc 2700

Qy   3221 tcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacaga 3280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 tcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacaga 2760

Qy   3281 accttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcca 3340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 accttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggcca 2820

Qy   3341 ccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaac 3400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 ccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaac 2880

Qy   3401 tggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagccca 3460
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2881 tggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcaaagccca 2940

Qy   3461 gcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagga 3520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 gcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctgcagga 3000

Qy   3521 gagattgcaacaccatcccacactgtccaggcctt 3555
          |||||||||||||||||||||||||||||||||||
Db   3001 gagattgcaacaccatcccacactgtccaggcctt 3035
```

RESULT 6
AAV05886
ID AAV05886 standard; DNA; 4039 BP.
XX
AC AAV05886;
XX
DT 01-JUN-1998 (first entry)
XX
DE Human Tbc-1 gene.
XX
KW tre-2; BUB2; cdc16; mast cell; probe; hybridisation; subtraction method;
KW mouse; transcription factor; differentiation; proliferation; human; ds;
KW acute myelogenous leukaemia.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 403..3828
FT /*tag= a
FT /product= Tbc1_protein
XX
PN US5700927-A.
XX
PD 23-DEC-1997.

XX
PF 23-DEC-1994; 94US-0363300.
XX
PR 23-DEC-1994; 94US-0363300.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Richardson P, Zon L;
XX
DR WPI; 1998-062437/06.
DR P-PSDB; AAW44777.
XX
PT DNA encoding Tbc1 polypeptide - useful for treating leukaemia
XX
PS Claim 3; Fig 1A-B; 22pp; English.
XX
CC This is the nucleotide sequence which encodes a Tbc1 (tre-2, BUB2 and
CC cdc16) polypeptide, isolated from a bone marrow-derived mast cell DNA
CC library. The screen was carried out using a probe generated by a
CC subtraction method which compared mRNA expression in an undifferentiated
CC mast cell line P815 and cell line PGT6 (P815 cells transformed to express
CC the murine GATA-1 transcription factor - a factor which controls the
CC expression of genes involved in mast cell differentiation). Tbc1 encodes
CC a protein involved in the coupling of cell proliferation to cell
CC differentiation, which can be used to treat leukaemia (especially acute
CC myelogenous leukaemia) by causing leukaemic cells to differentiate.
XX
SQ Sequence 4039 BP; 1042 A; 1096 C; 1060 G; 841 T; 0 other;

Query Match 70.6%; Score 2510.4; DB 19; Length 4039;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 2949; Conservative 0; Mismatches 536; Indels 37; Gaps 8;

```
Qy      1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59
          |||||  |||| |||||||||| ||||| |||| | || |||||||| ||||| |||||
Db    292 atggaggcaatcacattcacagcgaggaagcatccgtttccctaacgaagtctccgtgga 351

Qy     60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
           || ||| ||||||||||  |||||| | || |||||||| || ||||| |||||||||
Db    352 cttcggcttgcagctggttgggtctccttacccgtgcattctctcaccactatgcccatg 411

Qy    118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
          ||||| ||||| || || ||||| ||||||||||  | |||| | | | |||||| |||
Db    412 ctgccgtgggtggtagccgaggtacgaagactcagtggccagtgctctaaaaaggagcct 471

Qy    178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
             || |||||||||| ||||| ||||| ||||| | | || || |||||  |
Db    472 aggacaaagcaagtccgcctttgggtttcgccctccggactgcggtgcgagcctgacctg 531

Qy    238 gggagaagtcaacagtgggatccccctgatctattccagcatctttgagtgcaagcctcag 297
          | || ||||||||  ||||| || || |||| ||||||||||||||||||||||||||||
Db    532 gagaaaagtcaaccatgggacccgctcatctgttccagcatctttgagtgcaagcctcag 591

Qy    298 cgtgttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
          ||||| |||||||||||||||||||||| ||||| || |||||||||| || ||||||
Db    592 cgtgtgcacaaactgattcacaacagtcacgacccgagctactttgcttgcctcattaag 651

Qy    358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
          || || |||| ||| |||||||| |||||||||||||| ||||| || ||||||||||||
Db    652 gaggatgctgcgcacaggcagagcctctgctatgtgtttaaagcagacgatcaaacaaaa 711

Qy    418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
          |||||||||||||||||||||||||| ||||| |||||||| |||||  ||||||
Db    712 gtgcctgagatcatcagctccatccggcaggccgggaagattgcccgccaggaagagctg 771

Qy    478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
          |  ||||| |||||||||||||| || ||  |||| ||||||||||||||||||| |||
Db    772 cgttgcccctccgagttcgacgataccttcgccaaaaagttcgaggtgctcttctgtggc 831

Qy    538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
          || ||||| ||||| |||||||||| || || || ||||| |||||||| |||||||||
Db    832 cgggtgactgtggctcacaagaaggccccacccgcactgattgacgagtgtatcgagaag 891

Qy    598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
          ||||| || || ||| |  | ||  |  | ||  |       |||| |  ||  |   |
Db    892 ttcaaccatgtgagctgtggtcgcagaacggactgggaagcgcccaccgggcagccatca 951

Qy    658 gcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagccc 717
          ||                |||||  | ||||||||||||||| ||||||||  |||||
Db    952 gc------------------gcctggccccaggcccatgcgcaaatccttctcacagcct 993

Qy    718 ggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggc 777
          || |||||||||||||||| ||||||||| | ||||| |   ||||||| || |||  |
Db    994 ggactgcgctcgctggccttcaggaaggagttccaggacgctagcctccgcagtagcacc 1053

Qy    778 ttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatatt 837
          ||         || |   || |  ||||| ||||||||||||||  | || ||||||  |
Db   1054 tt------tagctcctttgacaatgacatagagaaccacctcatcggtgggcacaatgtg 1107

Qy    838 gtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaa 897
          || |||||||||| || ||||| || |||||||||| ||||||||||||  ||||||
Db   1108 gttcagcccacagacatggaggagaaccgaactatgctgttcacgattggcccatctgaa 1167

Qy    898 gtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaatttttaaggagata 957
          |||||||||||||||||||||||||||||| || ||| ||||||||||||||||||||||
Db   1168 gtttacctcatcagtcctgacaccaaaaagattgcactggagaaaaatttttaaggagata 1227

Qy    958 tccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtct 1017
          |||||||||||||||||||||||||||||| |||||||||||||| || ||||| || ||||
Db   1228 tccttttgctctcagggcatcagacatgtggaccactttggattcatctgccgagagtgc 1287

Qy   1018 tc------cggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
          ||      |||  ||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1288 tcgggtggcggcagtggcggctttcattttgtctgttacgtgttccagtgcacaaatgaa 1347

Qy   1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
          |||||||||||  || || |||||||| ||||| ||||| |||||||| || || ||||||
Db   1348 gctctggttgacgagatcatgatgactctgaagcaggctttcacggtagctgcggtgcag 1407

Qy   1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
          ||||| |||||||| |||||||||| |||||||||||||||| ||||| ||||||||||
Db   1408 cagacggctaaggcaccagcccagctctgtgagggctgccccttgcaaggcctgcacaag 1467

Qy   1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
          ||||| || |||||||||||||||||||| || ||||| ||| ||||||| || ||||||
Db   1468 ctctgcgaaaggatagagggaatgaattcatctaaaaccaaattagaactccagaagcac 1527

Qy   1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
            |||| ||| | |||||||||||||||||| || || || || ||||||||||||||||||
Db   1528 ttgaccacactgaccaatcaggagcaggccaccatattcgaggaggttcagaaattgaga 1587

Qy   1312 ccgagaaatgagcagcgagagaatgaattgattatttcttttctgagatgtttatatgaa 1371
          || ||||| |||||||||||||||||||||| |||||||||||||||| || |||||||||
Db   1588 ccaagaaacgagcagcgagagaatgaattaattatttcttttctgaggtgcttatatgaa 1647

Qy   1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
          ||||| || ||||| |||| ||| | ||||| |   |||||||||   ||| ||||||||
Db   1648 gagaagcaaaaagagcacagccacactggggcgccaaagcagacactacaggtggcagca 1707

Qy   1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
          |||||||||| |||||  | |||||||||| |  || || ||| |||||  ||||||
Db   1708 gagaatattgggagtgacctgccacccagtgctagccggttcaggttagattcgctgaag 1767

Qy   1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1551
          ||||  |||||||||| || ||||||||||  |||| || ||| ||||||||||||||||
Db   1768 aacagagcaaagaggtccttaacagagtccctagagagcattctgtcccggggtaataaa 1827

Qy   1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
          |||||||||||||||| || |||  |||||||||||||| ||||||   ||||||||
Db   1828 gccagaggcctgcaggaccattccgccagtgtggatctggacagctccacttctagtact 1887

Qy   1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
            ||||||||||||||||||||  ||  || |||| |||||||||| ||| ||||||||
Db   1888 ctaagtaacaccagcaaagagctgtccatgggtgacaaggaggccttccccgtctctgag 1947
```

```
Qy   1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
Db   1948 acctccttcaagctccttggctcctcagatgacctgtccagtgactcagagggccacatt 2007

Qy   1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
Db   2008 gcagaagagtctgccctgttgtcaccccagcaggcgttcagaaggagagccaacaccctg 2067

Qy   1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
Db   2068 agtcatttcccagtagagtgccctgcgcctccagaacctgcccagagctctccaggggtc 2127

Qy   1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
Db   2128 tctcaaaggaaactcatgcggtaccactccgtgagcacagagacgcctcatgaacgcaag 2187

Qy   1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
Db   2188 gactttgaatccaaagcaaaccacctgggtgacacagatgggacccccgtgaagacccgg 2247

Qy   1972 aggcattcctggaggcagcagatattcctccgagtagccacccccgcagaaggcgtgcgat 2031
Db   2248 cggcactcgtggagacagcagatattccttcgagtggccactccacagaaggcttgtgac 2307

Qy   2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
Db   2308 tccccgagcagatatgaagattattccgagctgggagagctccctccacgctcccctttta 2367

Qy   2092 gaaccagtttgtgaagatgggccctttggcccccaccagaggaaaagaaaaggacatct 2151
Db   2368 gaaccggtgtgtgaggacggcccatttgg-ccagtacaggaagaaaagaggaagacgtca 2426

Qy   2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg-ctgcttagaat 2210
Db   2427 cgcgagcttcgagagctgtggaaaaaggccatcttgcagcagatcctgcctcgtcaggat 2486

Qy   2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
Db   2487 ggagaaggagaatcagaagctacaagcctctgaaaacgatttgctgaacaaacgcctcaa 2546

Qy   2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
Db   2547 gcttgactatgaagaaatcactccgtgtcttaaagaagtcactacagtgtgggaaaagat 2606

Qy   2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
Db   2607 gcttagcactccaggaagatccaaaattaagtttgacatggaaaaagtgcactcagctgt 2666

Qy   2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
Db   2667 tgggcaaggtgtgccacgtcatcaccgaggtgagatctggaaatttctagctgagcagtt 2726

Qy   2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
Db   2727 ccaccttaaacacccatttcctagtaaacagcagccaaaggacgtgccctacaaagagct 2786

Qy   2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
Db   2787 cctgaagaagctgacctcgcagcagcacgccattctcatcgacctcgggcgaacctttcc 2846

Qy   2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
Db   2847 aacacatccatacttctctgcccagcttggagcaggtcagctgtcactttacaacattct 2906

Qy   2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
Db   2907 gaaggcctactcgcttctggaccaggaggttggatactgccaaggtctcagctttgtggc 2966

Qy   2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
Db   2967 aggcattttgcttcttcacatgagtgaggaagaggcgttcaagatgctcaagttcctgat 3026

Qy   2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
Db   3027 gtttgacatggggctgcggaaacagtatcggccagacatgattattttgcagatccagat 3086

Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
Db   3087 gtaccagctgtcacggctcctccacgattaccaccgagacctctacaaccacctggaaga 3146

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
Db   3147 gcacgagactggccccctacgtacgcggctccctggtttctcaccgtgttcgcctcaca 3206

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
Db   3207 gttcccactcggctttgtagccagagtctttgatatgatcttccttcagggatcagaggt 3266

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
Db   3267 catatttaaagtagctttaagtcttttggggagccataagcccttgattctacagcatga 3326

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
Db   3327 gaacctggaaaccatcgtggacttcataaagaacacactccccaacctgggcctggtgca 3386

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
Db   3387 gatggagaagaccatcagtcaggtgtttgagatggacatcgccaagcagctccaggccta 3446

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
Db   3447 tgaggtcgagtaccacgtcgtccaggaggagcttattgagtcctcgcctctcagtgacaa 3506

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
Db   3507 ccaaagaatggagaaattggagaaaaccaacagcacgttgcgcaaacagaaccttgacct 3566

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
Db   3567 cctggagcagttgcaggtggcaaatgctaggatccaaagccttgaagccacggtagagaa 3626

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
Db   3627 acttcttaccagcgagagtaagctgaagcagcgtgcgctgaccctggaggtggagcg--t 3684

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
Db   3685 cgccctgctgcagatggtggaggagctgcggaggcaaagcgcccggcccagcactccaga 3744

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
Db   3745 gccagactgcacccagctggagcccacaggcgattgaccgct 3786
```

```
RESULT   7
AAH18401
ID   AAH18401 standard; cDNA; 2362 BP.
XX
AC   AAH18401;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:18465.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
```

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18465; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2362 BP; 657 A; 581 C; 578 G; 546 T; 0 other;

Query Match 59.4%; Score 2112; DB 22; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1444 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 1503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 60

Qy   1504 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 1563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 120

Qy   1564 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 1623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 180

Qy   1624 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttttaag 1683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttttaag 240

Qy   1684 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 1743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 300

Qy   1744 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 1803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 360

Qy   1804 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 1863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 420

Qy   1864 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 1923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 480

Qy   1924 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 1983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 540

Qy   1984 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 2043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 600

Qy   2044 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 2103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 660

Qy   2104 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 2163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 720

Qy   2164 gagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaat 2223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaat 780

Qy   2224 cagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaa 2283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 cagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaa 840

Qy   2284 gaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactcca 2343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactcca 900

Qy   2344 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 2403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 960

Qy   2404 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacac 2463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacac 1020

Qy   2464 cagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctg 2523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 cagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctg 1080

Qy   2524 acttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatac 2583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 acttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatac 1140

Qy   2584 ttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactca 2643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactca 1200

Qy   2644 cttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgctt 2703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 cttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgctt 1260

Qy   2704 cttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatgggg 2763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 cttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatgggg 1320

Qy   2764 ctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 2823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 1380

Qy   2824 aggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggc 2883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 aggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggc 1440

Qy   2884 cccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctggga 2943
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 cccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctggga 1500

Qy    2944 ttcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtg 3003
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 ttcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtg 1560

Qy    3004 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 3063
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 1620

Qy    3064 atagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaagacc 3123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 atagttgactttataaaaagcacgctacccaaccttggcttggtacagatggaaaagacc 1680

Qy    3124 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtac 3183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtac 1740

Qy    3184 cacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggat 3243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 cacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggat 1800

Qy    3244 aaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttg 3303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 aaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttg 1860

Qy    3304 caggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagcagt 3363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 caggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagcagt 1920

Qy    3364 gagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcag 3423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 gagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcag 1980

Qy    3424 acggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacg 3483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 acggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacg 2040

Qy    3484 cagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatcccacac 3543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 cagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatcccacac 2100

Qy    3544 tgtccaggcctt 3555
           ||||||||||||
Db    2101 tgtccaggcctt 2112


RESULT   8
AAS34787
ID   AAS34787 standard; cDNA; 1422 BP.
XX
AC   AAS34787;
XX
DT   04-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human neoplastic disease associated polypeptide #21.
XX
KW   Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW   hyperproliferative disorder; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155163-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01358.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
```

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465558/50.
DR P-PSDB; AAU21588.
XX
PT Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid
PT arthritis -
XX
PS Claim 4; SEQ ID No 31; 687pp; English.
XX
CC The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of disorders involving neoplastic disease such as
CC hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC also be useful for treating other disorders such as neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC and renal disorders. The polynucleotide sequences of the invention are
CC also useful in gene therapy. AAS34767-AAS35050 represent cDNA sequences
CC encoding for the novel human neoplastic disease associated polypeptides
CC of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1422 BP; 399 A; 322 C; 337 G; 364 T; 0 other;

Query Match 28.2%; Score 1000.8; DB 22; Length 1422;
Best Local Similarity 99.3%; Pred. No. 1.5e-262;
Matches 1005; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2544 tcttattgaccttgggcgaacctttcctacacacccatacttctctgcccagcttggagc 2603
Db 139 tattcttttcacagggcgaacctttcctacacacccatacttctctgcccagcttggagc 198

Qy 2604 aggacagctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtggg 2663
Db 199 aggacagctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtggg 258

Qy 2664 atattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaaga 2723
Db 259 atattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaaga 318

Qy 2724 ggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggcc 2783
Db 319 ggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggcc 378

Qy 2784 agacatgattattttacagatccagatgtaccagctctcgaggttgcttcatgattacca 2843
Db 379 agacatgattattttacagatccagatgtaccagctctcgaggttgcttcatgattacca 438

Qy 2844 cagagacctctacaatcacctggaggagcacgagatcggccccagcctctacgctgcccc 2903
Db 439 cagagacctctacaatcacctggaggagcacgagatcggccccagcctctacgctgcccc 498

Qy 2904 ctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccagagtctttga 2963
Db 499 ctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccagagtctttga 558

Qy 2964 tatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaag 3023
Db 559 tatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaag 618

Qy 3024 ccataagcccttgattctgcagcatgaaaacctagaaaccatagttgactttataaaaag 3083
Db 619 ccataagcccttgattctgcagcatgaaaacctagaaaccatagttgactttataaaaag 678

Qy 3084 cacgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaat 3143
Db 679 cacgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaat 738

Qy 3144 ggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaagaagaact 3203
Db 739 ggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaagaagaact 798

```
Qy   3204 tatcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacag 3263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 tatcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacag 858

Qy   3264 cagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggat 3323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 cagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggat 918

Qy   3324 ccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggc 3383
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 ccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggc 978

Qy   3384 catgcttaccttagaactggagcggtcggccctgctgcagacggtggaggagctgcggcg 3443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 catgcttaccttagaactggagcggtcggccctgctgcagacggtggaggagctgcggcg 1038

Qy   3444 gcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcga 3503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 gcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcga 1098

Qy   3504 ctgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 ctgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 1150


RESULT   9
AAK94796
ID   AAK94796 standard; cDNA; 2989 BP.
XX
AC   AAK94796;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human full-length cDNA, SEQ ID NO: 3914.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-0114089.
XX
PR   08-JUL-1999;   99JP-0194486.
PR   11-JAN-2000; 2000JP-0118774.
PR   02-MAY-2000; 2000JP-0183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
DR   WPI; 2001-524255/58.
DR   P-PSDB; AAM93840.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their
PT   use in genetic manipulation -
XX
PS   Claim 8; SEQ ID NO 3914; 1380pp + sequence listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been
CC   isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC   molecules have been determined. Primers for synthesising the full length
CC   cDNA are useful for clarifying the function of the protein encoded by
CC   the cDNA. The full length clones were obtained by construction of full
CC   length enriched cDNA libraries that were synthesised by the oligo-capping
CC   method. The primers enable the production of the full length cDNA easily
CC   without any special methods. The present sequence is a full length
CC   human cDNA of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ   Sequence 2989 BP; 795 A; 656 C; 698 G; 840 T; 0 other;


  Query Match             28.1%;  Score 999.4;  DB 22;  Length 2989;
  Best Local Similarity   99.9%;  Pred. No. 5.8e-262;
  Matches 1000;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2555 ttgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 2614
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 tggggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 257

Qy   2615 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 2674
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 317

Qy   2675 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 2734
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 377

Qy   2735 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 2794
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 437

Qy   2795 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 2854
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 497

Qy   2855 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 2914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 557

Qy   2915 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 2974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 617

Qy   2975 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 3034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 677

Qy   3035 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctaccca 3094
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    678 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctaccca 737

Qy   3095 accttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgcta 3154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    738 accttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgcta 797

Qy   3155 aacagttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcct 3214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    798 aacagttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcct 857

Qy   3215 ctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgca 3274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    858 ctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgca 917

Qy   3275 aacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttg 3334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    918 aacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttg 977

Qy   3335 aggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttacct 3394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978 aggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttacct 1037

Qy   3395 tagaactggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcag 3454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 tagaactggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcag 1097

Qy   3455 agcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctct 3514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 agcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctct 1157
```

```
Qy   3515 gcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          ||||||||||||||||||||||||||||||||||||||||||
Db   1158 gcaggagagattgcaacaccatcccacactgtccaggcctt 1198


RESULT  10
AAS34785
ID   AAS34785 standard; cDNA; 1269 BP.
XX
AC   AAS34785;
XX
DT   04-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human neoplastic disease associated polypeptide #19.
XX
KW   Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW   hyperproliferative disorder; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155163-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01358.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
```

```
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465558/50.
DR   P-PSDB; AAU21586.
XX
PT   Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT   diagnose diseases or disorders associated with aberrant expression or
PT   activity of polypeptides, and for treating cancers, rheumatoid
PT   arthritis  -
XX
PS   Claim 4; SEQ ID No 29; 687pp; English.
XX
CC   The present invention relates to the isolation of novel human neoplastic
CC   disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
CC   sequences encoding for these polypeptides. The sequences of the
CC   invention are useful in the diagnosis, treatment, prevention and/or
CC   prognosis of disorders involving neoplastic disease such as
CC   hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC   cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC   astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC   also be useful for treating other disorders such as neural disorders,
CC   immune system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC   and renal disorders. The polynucleotide sequences of the invention are
CC   also useful in gene therapy. AAS34767-AAS35050 represent cDNA sequences
CC   encoding for the novel human neoplastic disease associated polypeptides
CC   of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1269 BP; 362 A; 313 C; 277 G; 312 T; 5 other;

  Query Match             26.5%;  Score 943.2;  DB 22;  Length 1269;
  Best Local Similarity   98.7%;  Pred. No. 7.7e-247;
  Matches  967;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    2;

Qy   1908 aaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagac 1967
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 aagggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagac 351

Qy   1968 ccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtg 2027
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 ccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtg 411

Qy   2028 cgattcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctcc 2087
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    412 cgattcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctcc 471

Qy   2088 tttagaaccagtttgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggac 2147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 tttagaaccagtttgtgaagatgggccctttggcccccccaccagaggaaaagaaaaggac 531

Qy   2148 atctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttag 2207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    532 atctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttag 591

Qy   2208 aatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcct 2267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    592 aatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcct 651

Qy   2268 gaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaa 2327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    652 gaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaa 711

Qy   2328 gatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggc 2387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    712 gatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggc 771

Qy   2388 tgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagca 2447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    772 tgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagca 831

Qy   2448 attccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaaga 2507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    832 attccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaaga 891

Qy   2508 actcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaaccctt 2567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    892 actcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaaccctt 951

Qy   2568 tcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacat 2627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    952 tcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacat 1011

Qy   2628 tttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgt 2687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1012 tttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgt 1071

Qy   2688 agcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttct 2747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1072 agcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttct 1131

Qy   2748 gatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatcca 2807
           |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1132 gatgtttgacatggggctgcggaaacagtatcggccagac-tgattattttacagatcca 1190

Qy   2808 gatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctgga 2867
           |||||:||||||:||||:||||||||||||||||||||| | |||||| ||||||||||
Db   1191 gatgtwccagctytcgargttgcttcatgattaccacagana-ctctaccatcacctgga 1249

Qy   2868 ggagcacgagatcggcccca 2887
           || ||||| ||| | ||||
Db   1250 agaacacgaaatcnggccca 1269

RESULT  11
AAH04450
ID   AAH04450 standard; cDNA; 883 BP.
XX
AC   AAH04450;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:1285.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
```

PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 1285; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 883 BP; 245 A; 221 C; 219 G; 193 T; 5 other;

Query Match 22.1%; Score 785; DB 22; Length 883;
Best Local Similarity 96.7%; Pred. No. 9.9e-204;
Matches 843; Conservative 0; Mismatches 23; Indels 6; Gaps 4;

```
Qy   1444 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 1503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 60

Qy   1504 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 1563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 120

Qy   1564 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 1623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 180

Qy   1624 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttttaag 1683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctccttttaag 240

Qy   1684 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 1743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 300

Qy   1744 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 1803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 360

Qy   1804 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 1863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 420

Qy   1864 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 1923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 480

Qy   1924 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 1983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 540

Qy   1984 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 2043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 600

Qy   2044 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 2103
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    601 tatgaagattattcagagctgggagagctt-ccccacgatctcctttagaaccagtttgt 659

Qy   2104 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 2163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    660 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctncga 719

Qy   2164 gagc-tgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaa 2222
          |||| |||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db    720 aagcttgtggcaaaaggctattcttcaacagatactgntgcttaaaatggagaaggaaaa 779

Qy   2223 tcagaagctccaagcc---tctgaaaatgatttgctgaacaagcgcc-tgaagctcgatt 2278
          ||||||||| | |  |   |  |||  ||||  |||||||||||| |||| ||||||
Db    780 tcagaagcttccaagccttttgaaaatggattgcttgaacaagcgccttgaacttcgatt 839

Qy   2279 atgaagaaattactccctgtcttaaagaagta 2310
          ||||||||||||||  |  ||||||||||
Db    840 ntgaagaaattactcctggctttaaagaagta 871
```

RESULT 12
AAK92006
ID AAK92006 standard; cDNA; 849 BP.
XX
AC AAK92006;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA 5'-end sequence, SEQ ID NO: 466.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX

```
PT   830 Primers useful for synthesizing full length cDNA clones and their
PT   use in genetic manipulation -
XX
PS   Claim 2; SEQ ID NO 466; 1380pp + sequence listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been
CC   isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC   molecules have been determined. Primers for synthesising the full length
CC   cDNA are useful for clarifying the function of the protein encoded by
CC   the cDNA. The full length clones were obtained by construction of full
CC   length enriched cDNA libraries that were synthesised by the oligo-capping
CC   method. The primers enable the production of the full length cDNA easily
CC   without any special methods. The present sequence is the nucleotide
CC   sequence of the 5'-end of a cDNA provided in the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ   Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;


  Query Match             16.2%;  Score 574.2;  DB 22;  Length 849;
  Best Local Similarity   98.2%;  Pred. No. 3.5e-146;
  Matches  642;  Conservative    0;  Mismatches    6;  Indels    6;  Gaps    6;

Qy   2555 ttgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 2614
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 tggggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 257

Qy   2615 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 2674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 317

Qy   2675 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 2734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 377

Qy   2735 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 2794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 437

Qy   2795 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 2854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 497

Qy   2855 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 2914
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 557

Qy   2915 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 2974
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 617

Qy   2975 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 3034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 677

Qy   3035 tgattctgcagcatgaaaacctagaaaccatagttgacttttataaaaagcacgctaccca 3094
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    678 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgct-ccca 736

Qy   3095 accttggcttggtacagatggaaaagaccatcaatcaggta-tttgaaatggacatc-gc 3152
           |||||||||||| |||||||||||||||||||||||| ||| |||| |||||||||| |
Db    737 accttggcttgg-acagatggaaaagaccatcaatcangtattttgnaatggacatcggt 795

Qy   3153 taaacagttacaagcttatg-aagttgagtacca-cgtccttcaagaagaactt 3204
           |||||||||||||||||| |||||||| ||| ||||||| ||||||||||
Db    796 taaacagttacaagcttatgaaagttgagtnccaccgtcctttaagaagaactt 849


RESULT  13
AAK93666
ID   AAK93666 standard; cDNA; 849 BP.
XX
AC   AAK93666;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human cDNA clone representative sequence, SEQ ID NO: 2126.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-0114089.
XX
PR   08-JUL-1999;   99JP-0194486.
PR   11-JAN-2000; 2000JP-0118774.
PR   02-MAY-2000; 2000JP-0183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
DR   WPI; 2001-524255/58.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their
PT   use in genetic manipulation -
XX
PS   Example 11; SEQ ID NO 2126; 1380pp + sequence listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been
CC   isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC   molecules have been determined. Primers for synthesising the full length
CC   cDNA are useful for clarifying the function of the protein encoded by
CC   the cDNA. The full length clones were obtained by construction of full
CC   length enriched cDNA libraries that were synthesised by the oligo-capping
CC   method. The primers enable the production of the full length cDNA easily
CC   without any special methods. The present sequence was used as the
CC   representative sequence from a human clone which was used in
CC   homology searches to identify the clone.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ   Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;


  Query Match             16.2%;  Score 574.2;  DB 22;  Length 849;
  Best Local Similarity   98.2%;  Pred. No. 3.5e-146;
  Matches  642;  Conservative    0;  Mismatches    6;  Indels    6;  Gaps    6;

Qy   2555 ttgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 2614
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 tggggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctat 257

Qy   2615 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 2674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 cgctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaag 317

Qy   2675 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 2734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 gtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaa 377

Qy   2735 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 2794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 tgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgatta 437

Qy   2795 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 2854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 ttttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctct 497
```

```
Qy   2855 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 2914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498 acaatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctca 557

Qy   2915 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 2974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 ccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttc 617

Qy   2975 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 3034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618 ttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagccct 677

Qy   3035 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctaccca 3094
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    678 tgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgct-ccca 736

Qy   3095 accttggcttggtacagatggaaaagaccatcaatcaggta-tttgaaatggacatc-gc 3152
          |||||||||||| |||||||||||||||||||||||| ||| |||| |||||||||| |
Db    737 accttggcttgg-acagatggaaaagaccatcaatcangtattttgnaatggacatcggt 795

Qy   3153 taaacagttacaagcttatg-aagttgagtacca-cgtccttcaagaagaactt 3204
          |||||||||||||||||||| ||||||||| ||| ||||||| ||||||||||
Db    796 taaacagttacaagcttatgaaagttgagtnccaccgtcctttaagaagaactt 849


RESULT 14
AAZ50904
ID   AAZ50904 standard; DNA; 17590 BP.
XX
AC   AAZ50904;
XX
DT   31-MAY-2000  (first entry)
XX
DE   Human TBC-1 partial genomic DNA comprising 5' end sequence.
XX
KW   TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;
KW   Single nucleotide polymorphism; tissue differentiation; prostate cancer;
KW   linkage analysis; genetic map; detection; diagnosis; genotyping;
KW   transgenic animal; screening; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_signal      1..2000
FT                    /*tag= a
FT                    /note= "5' Regulatory region"
FT   exon             2001..2077
FT                    /*tag= b
FT                    /number= 1
FT   intron           2078..12739
FT                    /*tag= c
FT                    /number= 1
FT   misc_feature     9391..9845
FT                    /*tag= d
FT                    /note= "Amplicon 99-430"
FT   primer_bind      9391..9408
FT                    /*tag= e
FT                    /bound_moiety= "Primer B1"
FT                    /note= "Amplification of amplicon 99-430"
FT   primer_bind      9475..9493
FT                    /*tag= f
FT                    /bound_moiety= "Primer D1"
FT                    /note= "Microsequencing of marker 99-430-352"
FT   misc_binding     9482..9506
FT                    /*tag= g
FT                    /bound_moiety= "Probe P1"
FT                    /note= "Detection of Biallelic marker 99-430-352"
FT   misc_feature     9494
FT                    /*tag= h
FT                    /note= "Ambiguity base 'R' corresponds to 'A' in allele-1
FT                    and 'G' in allele-2 of biallelic marker 99-430-352"
FT   primer_bind      complement (9495..9513)
FT                    /*tag= i
FT                    /bound_moiety= "Primer E1"
FT                    /note= "Microsequencing of marker 99-430-352"
FT   primer_bind      complement (9828..9845)
FT                    /*tag= j
FT                    /bound_moiety= "Primer C1"
FT                    /note= "Amplification of amplicon 99-430"
FT   exon             12292..12373
FT                    /*tag= k
FT                    /number= 1 bis
FT   intron           12374..12739
FT                    /*tag= l
FT                    /number= 1 bis
FT   exon             12740..13249
FT                    /*tag= m
FT                    /number= 2
FT   intron           13250..17590
FT                    /*tag= n
FT                    /number= 2
XX
PN   WO200008209-A2.
XX
PD   17-FEB-2000.
XX
PF   06-AUG-1999;   99WO-IB01444.
XX
PR   07-AUG-1998;   98US-0095653.
XX
PA   (GEST ) GENSET.
XX
PI   Blumenfeld M,  Bougueleret L,  Chumakov I;
XX
DR   WPI; 2000-205736/18.
XX
PT   New isolated human TBC-1 nucleic acids, useful for developing products
PT   for the diagnosis and treatment of disorders involving cell
PT   proliferation, particularly prostate cancer  -
XX
PS   Claim 1; Page 93-100; 166pp; English.
XX
CC   The present sequence is the partial genomic DNA of human TBC-1 gene,
CC   comprising the 5' regulatory region, exons 1, 1bis and 2. TBC-1 gene is
CC   mapped to a candidate region of prostate cancer on chromosome 4. Single
CC   nucleotide polymorphism (SNP) is located within the biallelic marker
CC   region 99-430-352, localised in intron 1 of TBC-1 genomic DNA.
CC   TBC-1 gene is involved in the regulation of cell cycle and tissue
CC   differentiation in mammals. An alteration of TBC-1 sequence may be
CC   associated with a pathological condition, resulting in abnormal cell
CC   proliferation leading to cancer, e.g. prostate cancer. The biallelic
CC   markers can be used for generation of genetic maps, linkage analysis and
CC   association studies. TBC-1 sequence can be used for detection,
CC   diagnosis, genotyping, production of transgenic animals and screening
CC   of compounds for use in therapy.
XX
SQ   Sequence 17590 BP; 4760 A; 3776 C; 4104 G; 4919 T; 31 other;


  Query Match             11.9%;  Score 422.6;  DB 21;  Length 17590;
  Best Local Similarity   95.8%;  Pred. No. 5.4e-104;
  Matches  434;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   12833 atggaaccaataacattcacagcaaggaaacatctgctttctaacgaggtctcggtggat 12892

Qy      61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12893 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 12952

Qy     121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12953 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 13012
```

```
Qy    181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13013 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 13072

Qy    241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13073 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 13132

Qy    301 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13133 gttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaaggaa 13192

Qy    361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13193 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagta 13252

Qy    421 cctgagatcatcagctccatccgtcaggcgggg 453
          ||||||    | |   |   | ||| | ||
Db  13253 agtgagatggagatccaaaagactaaggtgtgg 13285
```

```
RESULT  15
AAZ15229
ID   AAZ15229 standard; cDNA; 696 BP.
XX
AC   AAZ15229;
XX
DT   12-OCT-1999  (first entry)
XX
DE   Human gene expression product cDNA sequence SEQ ID NO:2698.
XX
KW   Human; gene; gene expression product; diagnosis; therapy; probe;
KW   detection; mapping; tissue typing; profiling; forensic; cancer;
KW   genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9938972-A2.
XX
PD   05-AUG-1999.
XX
PF   28-JAN-1999;   99WO-US01619.
XX
PR   03-APR-1998;   98US-0030666.
PR   28-JAN-1998;   98US-0072910.
PR   24-FEB-1998;   98US-0075954.
PR   31-MAR-1998;   98US-0080114.
PR   03-APR-1998;   98US-0080515.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Crkvenjakov R,  Dickson M,  Drmanac R,  Drmanac S;
PI   Escobedo J,  Garcia PD,  Garcia V,  Giese K,  Innis MA;
PI   Jones WL,  Kassam A,  Kennedy GC,  Kita D,  Labat I;
PI   Lamson G,  Leshkowitz D,  Pot D,  Randazzo F,  Reinhard C;
PI   Stache-Crain B,  Sudduth-Klinger J,  Williams LT;
XX
DR   WPI; 1999-494092/41.
XX
PT   Novel human genes and their expression products which are
PT   differentially expressed in different cell types
XX
PS   Claim 1; Page 1311; 2479pp; English.
XX
CC   The present invention describes a library of human polynucleotides
CC   comprising the sequences given in AAZ12532 to AAZ17779. Also described is
CC   a method of detecting differentially expressed genes correlated with the
CC   cancerous state of a mammalian cell, comprising detecting at least one
CC   differentially expressed gene product in a test sample from a cell
CC   suspected of being cancerous, where the gene product is encoded by one
CC   of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC   polynucleotides can be used as a source of primers and probes, which can
CC   be used for a variety of purpose, e.g. detection of expression levels,
CC   mapping, tissue typing or profiling, forensics, genetic analysis and
CC   detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC   can be used for raising antibodies for experimental, diagnostic and
CC   therapeutic purposes. The polynucleotides may also be used to construct
CC   arrays for diagnostics (which may be used to determine function of an
CC   encoded protein); and to detect differences in expression levels between
CC   two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC   identify a genetic predisposition or susceptibility to a disease such as
CC   cancer). The polynucleotides of the invention are especially used in the
CC   diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC   and lung cancer. The polynucleotides can also be used to screen for
CC   peptide analogues and antagonists.
XX
SQ   Sequence 696 BP; 189 A; 157 C; 178 G; 142 T; 30 other;
```

```
  Query Match             10.2%;  Score 362;  DB 20;  Length 696;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-88;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   3194 aagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagaga 3253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     59 aagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagaga 118

Qy   3254 aaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaa 3313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 aaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaa 178

Qy   3314 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 3373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 238

Qy   3374 tgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggagg 3433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 tgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggagg 298

Qy   3434 agctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagc 3493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 agctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagc 358

Qy   3494 ccacgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcc 3553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 ccacgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcc 418

Qy   3554 tt 3555
          ||
Db    419 tt 420
```

```
Search completed: September  9, 2002, 19:18:41
Job time: 17499 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 14:24:47 ; Search time 8477.27 Seconds
(without alignments)
8775.683 Million cell updates/sec

Title: US-09-762-311-3_COPY_171_3725
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tcccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2751	77.4	3023	9	BC014529	BC014529 Homo sapi
2	2510.4	70.6	4039	6	I86429	I86429 Sequence 1
3	2510.4	70.6	4039	10	MMU33005	U33005 Mus musculu
4	2340	65.8	2576	9	AB029031	AB029031 Homo sapi
5	2112	59.4	2362	9	AK027355	AK027355 Homo sapi
6	1359.2	38.2	1995	4	BTA17923	Y17923 Bos taurus
7	1179.6	33.2	3051	10	BC004675	BC004675 Mus muscu
8	998.4	28.1	1788	9	AK057182	AK057182 Homo sapi
9	682.2	19.2	5922	9	AB011175	AB011175 Homo sapi
c 10	465.8	13.1	189811	30	AC009595	Ac009595 Homo sapi
11	465.8	13.1	195108	9	AC021106	AC021106 Homo sapi
c 12	465.8	13.1	208318	2	AC108933	AC108933 Homo sapi
c 13	424.2	11.9	110000	2	AL390202_04	Continuation (5 of
14	304.6	8.6	3214	3	DMU50542	U50542 Drosophila
15	304.6	8.6	4746	3	DME17919	Y17919 Drosophila
16	284.4	8.0	189811	30	AC009595	Ac009595 Homo sapi
17	284.4	8.0	208318	2	AC108933	AC108933 Homo sapi
18	252.4	7.1	163928	2	AC021335	AC021335 Homo sapi
19	252.4	7.1	180918	9	AC098680	AC098680 Homo sapi
20	245.6	6.9	17019	2	AC019886	AC019886 Drosophil
21	245.6	6.9	161117	3	AC008311	AC008311 Drosophil
22	245.6	6.9	168448	3	AC009346	AC009346 Drosophil
23	245.6	6.9	297832	3	AE003602	AE003602 Drosophil
c 24	180	5.1	163928	2	AC021335	AC021335 Homo sapi
25	141.8	4.0	3761	9	BC014111	BC014111 Homo sapi
c 26	141	4.0	221717	2	AC074357	AC074357 Mus muscu
c 27	139.2	3.9	158310	9	AL162571	AL162571 Human DNA
28	139.2	3.9	183577	2	AC011164	AC011164 Homo sapi
29	132.8	3.7	3691	6	AX078376	AX078376 Sequence
30	132.8	3.7	4560	9	HSA011679	AJ011679 Homo sapi
31	123.8	3.5	154	6	AX261661	AX261661 Sequence
32	119.4	3.4	14119	2	AC013898	AC013898 Drosophil
c 33	119.4	3.4	164413	3	AC023718	AC023718 Drosophil
c 34	119.4	3.4	303648	3	AE003486	AE003486 Drosophil
35	114.6	3.2	2013	9	AF042345	AF042345 Homo sapi
36	114.6	3.2	7383	9	AF008915	AF008915 Homo sapi
c 37	101.8	2.9	30027	3	CELT24D11	U49940 Caenorhabdi
c 38	98	2.8	20553	3	CBRG26M15	AC084540 Caenorhab
c 39	95.6	2.7	73803	8	NCB11B22	AL356834 Neurospor
40	90	2.5	5870	10	MMU53586	U53586 Mus musculu
c 41	84.2	2.4	441	11	G47994	G47994 Z27058_1 Ze
42	80	2.3	27102	3	U41540	U41540 Caenorhabdi
c 43	79.6	2.2	110032	2	AC014916	AC014916 Drosophil
44	79.6	2.2	159660	3	AC010040	AC010040 Drosophil
45	79.6	2.2	282267	3	AE003555	AE003555 Drosophil

ALIGNMENTS

RESULT 1
BC014529
LOCUS BC014529 3023 bp mRNA linear PRI 19-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:3843156, mRNA.
ACCESSION BC014529
VERSION BC014529.1 GI:17939551
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3023)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 14 Row: c Column: 16
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

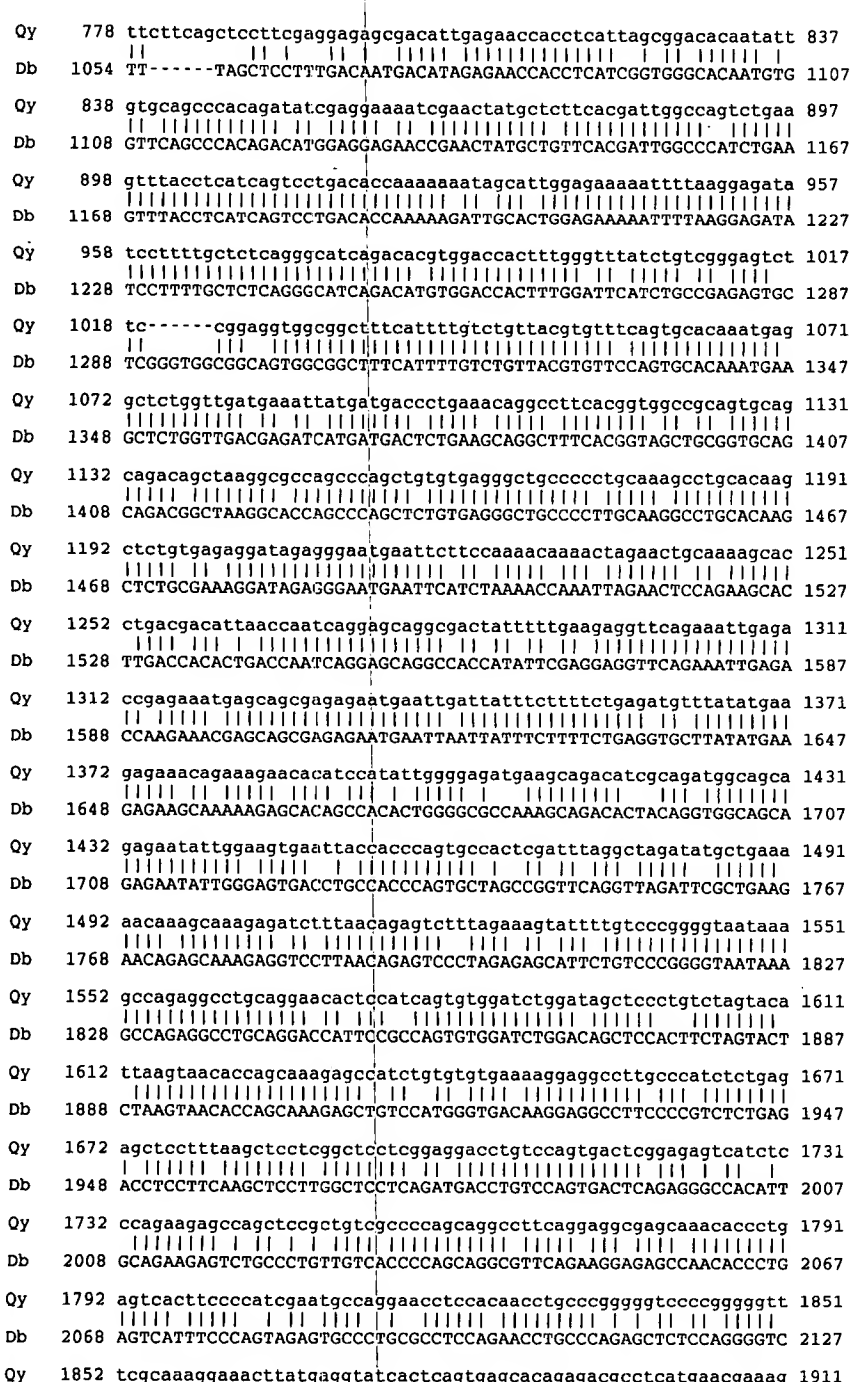
FEATURES Location/Qualifiers
source 1. .3023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3843156"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 883 a 712 c 740 g 688 t
ORIGIN

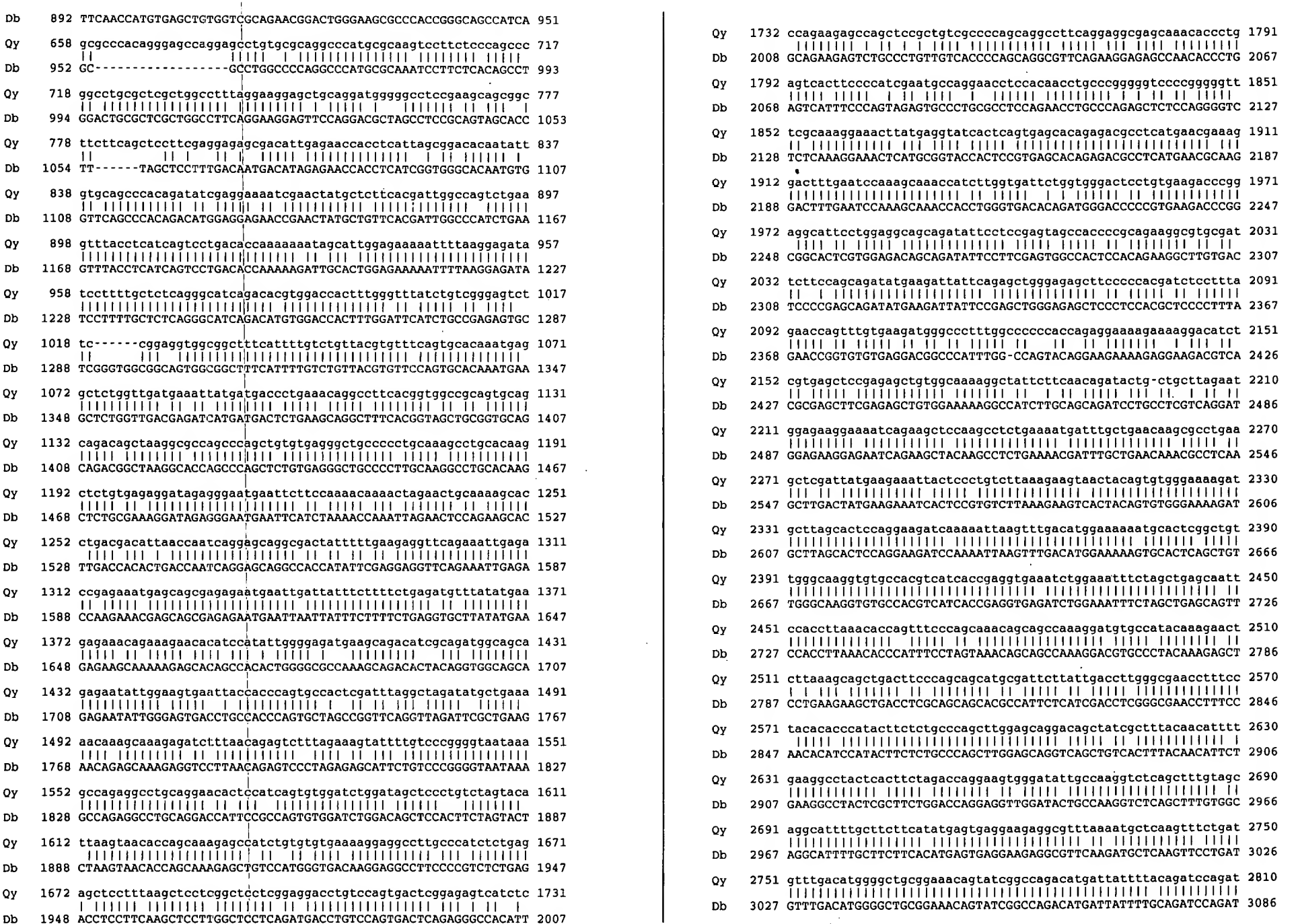
Query Match 77.4%; Score 2751; DB 9; Length 3023;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    805 attgagaaccacctcattagcggacacaatattgtgcagcccacagatatcgaggaaaat 864
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 ATTGAGAACCACCTCATTAGCGGACACAATATTGTGCAGCCCACAGATATCGAGGAAAAT 69

Qy    865 cgaactatgctcttcacgattggccagtctgaagtttacctcatcagtcctgacaccaaa 924
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 CGAACTATGCTCTTCACGATTGGCCAGTCTGAAGTTTACCTCATCAGTCCTGACACCAAA 129

Qy    925 aaaatagcattggagaaaaatttttaaggagatatccttttgctctcagggcatcagacac 984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 AAAATAGCATTGGAGAAAAATTTTAAGGAGATATCCTTTTGCTCTCAGGGCATCAGACAC 189

Qy    985 gtggaccactttgggtttatctgtcgggagtcttccggaggtggcggctttcattttgtc 1044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 GTGGACCACTTTGGGTTTATCTGTCGGGAGTCTTCCGGAGGTGGCGGCTTTCATTTTGTC 249

Qy   1045 tgttacgtgtttcagtgcacaaatgaggctctggttgatgaaattatgatgaccctgaaa 1104
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 TGTTACGTGTTTCAGTGCACAAATGAGGCTCTGGTTGATGAAATTATGATGACCCTGAAA 309

Qy   1105 caggccttcacggtggccgcagtgcagcagacagctaaggcgccagcccagctgtgtgag 1164
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 CAGGCCTTCACGGTGGCCGCAGTGCAGCAGACAGCTAAGGCGCCAGCCCAGCTGTGTGAG 369

Qy   1165 ggctgccccctgcaaagcctgcacaagctctgtgagaggatagagggaatgaattcttcc 1224
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 GGCTGCCCCCTGCAAAGCCTGCACAAGCTCTGTGAGAGGATAGAGGGAATGAATTCTTCC 429

Qy   1225 aaaacaaaactagaactgcaaaagcacctgacgacattaaccaatcaggagcaggcgact 1284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 AAAACAAAACTAGAACTGCAAAAGCACCTGACGACATTAACCAATCAGGAGCAGGCGACT 489

Qy   1285 atttttgaagaggttcagaaattgagaccgagaaatgagcagcgagagaatgaattgatt 1344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 ATTTTTGAAGAGGTTCAGAAATTGAGACCGAGAAATGAGCAGCGAGAGAATGAATTGATT 549

Qy   1345 atttcttttctgagatgtttatatgaagagaaacagaaagaacacatccatattggggag 1404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    550 ATTTCTTTTCTGAGATGTTTATATGAAGAGAAACAGAAAGAACACATCCATATTGGGGAG 609

Qy   1405 atgaagcagacatcgcagatggcagcagagaatattggaagtgaattaccacccagtgcc 1464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 ATGAAGCAGACATCGCAGATGGCAGCAGAGAATATTGGAAGTGAATTACCACCCAGTGCC 669

Qy   1465 actcgatttaggctagatatgctgaaaaacaaagcaaagagatctttaacagagtcttta 1524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    670 ACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAGAGATCTTTAACAGAGTCTTTA 729

Qy   1525 gaaagtatttttgtcccggggtaataaagccagaggcctgcaggaacactccatcagtgtg 1584
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    730 GAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTGCAGGAACACTCCATCAGTGTG 789

Qy   1585 gatctggatagctccctgtctagtacattaagtaacaccagcaaagagccatctgtgtgt 1644
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    790 GATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCATCTGTGTGT 849

Qy   1645 gaaaaggaggccttgcccatctctgagagctcctttaagctcctcggctcctcggaggac 1704
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    850 GAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCCTCGGAGGAC 909

Qy   1705 ctgtccagtgactcggagagtcatctcccagaagagccagctccgctgtcgccccagcag 1764
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    910 CTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCGCCCCAGCAG 969

Qy   1765 gccttcaggaggcgagcaaacaccctgagtcacttccccatcgaatgccaggaacctcca 1824
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    970 GCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAGGAACCTCCA 1029

Qy   1825 caacctgcccgggggtccccgggggtttcgcaaaggaaacttatgaggtatcactcagtg 1884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1030 CAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTATCACTCAGTG 1089

Qy   1885 agcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccatcttggtgat 1944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1090 AGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCATCTTGGTGAT 1149

Qy   1945 tctggtgggactcctgtgaagacccggaggcattcctggaggcagcagatattcctccga 2004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1150 TCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATATTCCTCCGA 1209

Qy   2005 gtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattattcagagctg 2064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1210 GTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTATTCAGAGCTG 1269

Qy   2065 ggagagcttcccccacgatctcctttagaaccagtttgtgaagatgggccctttggcccc 2124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1270 GGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCCTTTGGCCCC 1329

Qy   2125 ccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaaaaggctatt 2184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1330 CCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAAGGCTATT 1389

Qy   2185 cttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaagcctctgaa 2244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1390 CTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAAGCCTCTGAA 1449

Qy   2245 aatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccctgtcttaaa 2304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1450 AATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCCTGTCTTAAA 1509

Qy   2305 gaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaaattaagttt 2364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1510 GAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAAATTAAGTTT 1569

Qy   2365 gacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaa 2424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1570 GACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCATCACCGAGGTGAA 1629

Qy   2425 atctggaaatttctagctgagcaattccaccttaaacaccagtttcccagcaaacagcag 2484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1630 ATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGCAAACAGCAG 1689

Qy   2485 ccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcagcatgcgatt 2544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1690 CCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAGCATGCGATT 1749
```

```
Qy   2545 cttattgaccttgggcgaacctttcctacacacccatacttctctgcccagcttggagca 2604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1750 CTTATTGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAGCTTGGAGCA 1809

Qy   2605 ggacagctatcgctttacaacattttgaaggcctactcacttctagaccaggaagtggga 2664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1810 GGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAGACCAGGAAGTGGGA 1869

Qy   2665 tattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagag 2724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1870 TATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTTTGCTTCTTCATATGAGTGAGGAAGAG 1929

Qy   2725 gcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacagtatcggcca 2784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1930 GCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAAACAGTATCGGCCA 1989

Qy   2785 gacatgattatttttacagatccagatgtaccagctctcgaggttgcttcatgattaccac 2844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1990 GACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCTTCATGATTACCAC 2049

Qy   2845 agagacctctacaatcacctggaggagcacgagatcggccccagcctctacgctgccccc 2904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2050 AGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCTCTACGCTGCCCCC 2109

Qy   2905 tggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccagagtctttgat 2964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2110 TGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCAGAGTCTTTGAT 2169

Qy   2965 atgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagc 3024
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2170 ATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTCTGTTGGGAAGC 2229

Qy   3025 cataagcccttgattctgcagcatgaaaacctagaaaccatagttgactttataaaaagc 3084
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2230 CATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACTTTATAAAAAGC 2289

Qy   3085 acgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatg 3144
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2290 ACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGGTATTTGAAATG 2349

Qy   3145 gacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaagaagaactt 3204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2350 GACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTCAAGAAGAACTT 2409

Qy   3205 atcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagc 3264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2410 ATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGAAAACCAACAGC 2469

Qy   3265 agcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatc 3324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2470 AGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAAATGGTAGGATC 2529

Qy   3325 caaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggcc 3384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2530 CAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGCTGAAGCAGGCC 2589

Qy   3385 atgcttaccttagaactggagcggtcggccctgctgcagacggtggaggagctgcggcgg 3444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2590 ATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGGAGCTGCGGCGG 2649

Qy   3445 cggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgac 3504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2650 CGGAGCGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGCCCACGGGCGAC 2709

Qy   3505 tgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2710 TGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCCTT 2760


RESULT   2
I86429
LOCUS       I86429                 4039 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 1 from patent US 5700927.
ACCESSION   I86429
VERSION     I86429.1  GI:3206147
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4039)
  AUTHORS   Zon,L. and Richardson,P.
  TITLE     Tbc1 gene and uses thereof
  JOURNAL   Patent: US 5700927-A 1 23-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .4039
                     /organism="unknown"
BASE COUNT     1042 a   1096 c   1060 g    841 t
ORIGIN


  Query Match             70.6%;  Score 2510.4;  DB 6;  Length 4039;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches 2949;  Conservative    0;  Mismatches  536;  Indels   37;  Gaps    8;

Qy      1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59
          |||||  |||| ||||||||||| ||||| |||| | || |||||||| ||||| |||||
Db    292 ATGGAGGCAATCACATTCACAGCGAGGAAGCATCCGTTTCCCTAACGAAGTCTCCGTGGA 351

Qy     60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
           || ||| ||||||||||  || |||| | || |||||||| || ||||| |||||||||
Db    352 CTTCGGCTTGCAGCTGGTTGGGTCTCCTTACCCGTGCATTCTCTCACCACTATGCCCATG 411

Qy    118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
          ||||| |||| || || |||||  |||||||||||  | |||| | | | |||||| |||
Db    412 CTGCCGTGGGTGGTAGCCGAGGTACGAAGACTCAGTGGCCAGTGCTCTAAAAAGGAGCCT 471

Qy    178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
            || |||||||||| ||||| ||||| ||||| |||||| | || || ||||| |
Db    472 AGGACAAAGCAAGTCCGCCTTTGGGTTTCGCCCTCCGGACTGCGGTGCGAGCCTGACCTG 531

Qy    238 gggagaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcag 297
          | || ||||||||  ||||| || || |||| ||||||||||||||||||||||||||||
Db    532 GAGAAAAGTCAACCATGGGACCCGCTCATCTGTTCCAGCATCTTTGAGTGCAAGCCTCAG 591

Qy    298 cgtgttcacaaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
          ||||| |||||||||||||||||||||||| ||||| || |||||||||| || ||||||
Db    592 CGTGTGCACAAACTGATTCACAACAGTCACGACCCGAGCTACTTTGCTTGCCTCATTAAG 651

Qy    358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
          || || ||||  ||| |||||||  |||||||||||||| ||||| || ||||||||||||
Db    652 GAGGATGCTGCGCACAGGCAGAGCCTCTGCTATGTGTTTAAAGCAGACGATCAAACAAAA 711

Qy    418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
          |||||||||||||||||||||||||||| ||||| |||||||| ||||| ||||| ||||||
Db    712 GTGCCTGAGATCATCAGCTCCATCCGGCAGGCCGGGAAGATTGCCCGCCAGGAAGAGCTG 771

Qy    478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
          |  ||||| |||||||||||||| || ||  |||| ||||||||||||||||||| |||
Db    772 CGTTGCCCCTCCGAGTTCGACGATACCTTCGCCAAAAAGTTCGAGGTGCTCTTCTGTGGC 831

Qy    538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
          || ||||| |||||||||||| || || || ||||| |||||||| |||||||||
Db    832 CGGGTGACTGTGGCTCACAAGAAGGCCCCACCCGCACTGATTGACGAGTGTATCGAGAAG 891

Qy    598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
          ||||| || || ||| |  | || |  | ||  |       |||| |  ||  |   |
Db    892 TTCAACCATGTGAGCTGTGGTCGCAGAACGGACTGGGAAGCGCCCACCGGGCAGCCATCA 951

Qy    658 gcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagccc 717
          ||                    |||||  | |||||||||||||||| |||||||| |||||
Db    952 GC------------------GCCTGGCCCCAGGCCCATGCGCAAATCCTTCTCACAGCCT 993

Qy    718 ggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcggc 777
          || |||||||||||||||| ||||||||  | ||||| |   |||||| || |||  |
Db    994 GGACTGCGCTCGCTGGCCTTCAGGAAGGAGTTCCAGGACGCTAGCCTCCGCAGTAGCACC 1053
```


Qy 778 ttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaatatt 837
Db 1054 TT------TAGCTCCTTTGACAATGACATAGAGAACCACCTCATCGGTGGGCACAATGTG 1107
Qy 838 gtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgattggccagtctgaa 897
Db 1108 GTTCAGCCCACAGACATGGAGGAGAACCGAACTATGCTGTTCACGATTGGCCCATCTGAA 1167
Qy 898 gtttacctcatcagtcctgacaccaaaaaaatagcattggagaaaaattttaaggagata 957
Db 1168 GTTTACCTCATCAGTCCTGACACCAAAAAGATTGCACTGGAGAAAAATTTTAAGGAGATA 1227
Qy 958 tccttttgctctcagggcatcagacacgtggaccactttgggtttatctgtcgggagtct 1017
Db 1228 TCCTTTTGCTCTCAGGGCATCAGACATGTGGACCACTTTGGATTCATCTGCCGAGAGTGC 1287
Qy 1018 tc------cggaggtggcggctttcattttgtctgttacgtgtttcagtgcacaaatgag 1071
Db 1288 TCGGGTGGCGGCAGTGGCGGCTTTCATTTTGTCTGTTACGTGTTCCAGTGCACAAATGAA 1347
Qy 1072 gctctggttgatgaaattatgatgaccctgaaacaggccttcacggtggccgcagtgcag 1131
Db 1348 GCTCTGGTTGACGAGATCATGATGACTCTGAAGCAGGCTTTCACGGTAGCTGCGGTGCAG 1407
Qy 1132 cagacagctaaggcgccagcccagctgtgtgagggctgccccctgcaaagcctgcacaag 1191
Db 1408 CAGACGGCTAAGGCACCAGCCCAGCTCTGTGAGGGCTGCCCCTTGCAAGGCCTGCACAAG 1467
Qy 1192 ctctgtgagaggatagagggaatgaattcttccaaaacaaaactagaactgcaaaagcac 1251
Db 1468 CTCTGCGAAAGGATAGAGGGAATGAATTCATCTAAAACCAAATTAGAACTCCAGAAGCAC 1527
Qy 1252 ctgacgacattaaccaatcaggagcaggcgactatttttgaagaggttcagaaattgaga 1311
Db 1528 TTGACCACACTGACCAATCAGGAGCAGGCCACCATATTCGAGGAGGTTCAGAAATTGAGA 1587
Qy 1312 ccgagaaatgagcagcgagagaatgaattgattatttctttttctgagatgtttatatgaa 1371
Db 1588 CCAAGAAACGAGCAGCGAGAGAATGAATTAATTATTTCTTTTCTGAGGTGCTTATATGAA 1647
Qy 1372 gagaaacagaaagaacacatccatattggggagatgaagcagacatcgcagatggcagca 1431
Db 1648 GAGAAGCAAAAAGAGCACAGCCACACTGGGGCGCCAAAGCAGACACTACAGGTGGCAGCA 1707
Qy 1432 gagaatattggaagtgaattaccacccagtgccactcgatttaggctagatatgctgaaa 1491
Db 1708 GAGAATATTGGGAGTGACCTGCCACCCAGTGCTAGCCGGTTCAGGTTAGATTCGCTGAAG 1767
Qy 1492 aacaaagcaaagagatctttaacagagtctttagaaagtattttgtcccggggtaataaa 1551
Db 1768 AACAGAGCAAAGAGGTCCTTAACAGAGTCCCTAGAGAGCATTCTGTCCCGGGGTAATAAA 1827
Qy 1552 gccagaggcctgcaggaacactccatcagtgtggatctggatagctccctgtctagtaca 1611
Db 1828 GCCAGAGGCCTGCAGGACCATTCCGCCAGTGTGGATCTGGACAGCTCCACTTCTAGTACT 1887
Qy 1612 ttaagtaacaccagcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgag 1671
Db 1888 CTAAGTAACACCAGCAAAGAGCTGTCCATGGGTGACAAGGAGGCCTTCCCCGTCTCTGAG 1947
Qy 1672 agctcctttaagctcctcggctcctcggaggacctgtccagtgactcggagagtcatctc 1731
Db 1948 ACCTCCTTCAAGCTCCTTGGCTCCTCAGATGACCTGTCCAGTGACTCAGAGGGCCACATT 2007
Qy 1732 ccagaagagccagctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctg 1791
Db 2008 GCAGAAGAGTCTGCCCTGTTGTCACCCCAGCAGGCGTTCAGAAGGAGAGCCAACACCCTG 2067
Qy 1792 agtcacttccccatcgaatgccaggaacctccacaacctgcccgggggtccccgggggtt 1851
Db 2068 AGTCATTTCCCAGTAGAGTGCCCTGCGCCTCCAGAACCTGCCCAGAGCTCTCCAGGGGTC 2127
Qy 1852 tcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaag 1911
Db 2128 TCTCAAAGGAAACTCATGCGGTACCACTCCGTGAGCACAGAGACGCCTCATGAACGCAAG 2187
Qy 1912 gactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtgaagacccgg 1971
Db 2188 GACTTTGAATCCAAAGCAAACCACCTGGGTGACACAGATGGGACCCCCGTGAAGACCCGG 2247
Qy 1972 aggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgat 2031
Db 2248 CGGCACTCGTGGAGACAGCAGATATTCCTTCGAGTGGCCACTCCACAGAAGGCTTGTGAC 2307
Qy 2032 tcttccagcagatatgaagattattcagagctgggagagcttcccccacgatctccttta 2091
Db 2308 TCCCCGAGCAGATATGAAGATTATTCCGAGCTGGGAGAGCTCCCTCCACGCTCCCCTTTA 2367
Qy 2092 gaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaaaggacatct 2151
Db 2368 GAACCGGTGTGTGAGGACGGCCCATTTGG-CCAGTACAGGAAGAAAAGAGGAAGACGTCA 2426
Qy 2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactg-ctgcttagaat 2210
Db 2427 CGCGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTGCAGCAGATCCTGCCTCGTCAGGAT 2486
Qy 2211 ggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaa 2270
Db 2487 GGAGAAGGAGAATCAGAAGCTACAAGCCTCTGAAAACGATTTGCTGAACAAACGCCTCAA 2546
Qy 2271 gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagat 2330
Db 2547 GCTTGACTATGAAGAAATCACTCCGTGTCTTAAAGAAGTCACTACAGTGTGGGAAAAGAT 2606
Qy 2331 gcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgt 2390
Db 2607 GCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAAGTGCACTCAGCTGT 2666
Qy 2391 tgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaatt 2450
Db 2667 TGGGCAAGGTGTGCCACGTCATCACCGAGGTGAGATCTGGAAATTTCTAGCTGAGCAGTT 2726
Qy 2451 ccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatacaaagaact 2510
Db 2727 CCACCTTAAACACCCATTTCCTAGTAAACAGCAGCCAAAGGACGTGCCCTACAAAGAGCT 2786
Qy 2511 cttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctttcc 2570
Db 2787 CCTGAAGAAGCTGACCTCGCAGCAGCACGCCATTCTCATCGACCTCGGGCGAACCTTTCC 2846
Qy 2571 tacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaacatttt 2630
Db 2847 AACACATCCATACTTCTCTGCCCAGCTTGGAGCAGGTCAGCTGTCACTTTACAACATTCT 2906
Qy 2631 gaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagctttgtagc 2690
Db 2907 GAAGGCCTACTCGCTTCTGGACCAGGAGGTTGGATACTGCCAAGGTCTCAGCTTTGTGGC 2966
Qy 2691 aggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgat 2750
Db 2967 AGGCATTTTGCTTCTTCACATGAGTGAGGAAGAGGCGTTCAAGATGCTCAAGTTCCTGAT 3026
Qy 2751 gtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatccagat 2810
Db 3027 GTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTGCAGATCCAGAT 3086
Qy 2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
Db 3087 GTACCAGCTGTCACGGCTCCTCCACGATTACCACCGAGACCTCTACAACCACCTGGAAGA 3146
Qy 2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
Db 3147 GCACGAGACTGGCCCCCCTACGTACGCGGCTCCCTGGTTTCTCACCGTGTTCGCCTCACA 3206
Qy 2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990

```
Db    3207 GTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGATCAGAGGT 3266

Qy    2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
           |||||||||||| |||||||||| ||||| ||||||||||||||||||| ||||||||
Db    3267 CATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTACAGCATGA 3326

Qy    3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
           ||||| |||||||| || |||| |||| | ||| || |||||||| ||| |||| ||
Db    3327 GAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGCCTGGTGCA 3386

Qy    3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
           |||||| |||||||||| ||||||| ||||| |||||||||| || ||| | || || ||
Db    3387 GATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTCCAGGCCTA 3446

Qy    3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
           ||| || |||||||||||| || || || || ||||| || ||||| ||||||||||||||
Db    3447 TGAGGTCGAGTACCACGTCGTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTCAGTGACAA 3506

Qy    3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
           |||||||||||| ||||| ||||||||||||||||  || ||||||||||||||||||||
Db    3507 CCAAAGAATGGAGAAATTGGAGAAAACCAACAGCACGTTGCGCAAACAGAACCTTGACCT 3566

Qy    3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
           ||| || |||||||||||||||||| |||||||||||||||| ||||| | ||||||
Db    3567 CCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACGGTAGAGAA 3626

Qy    3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
           || || | ||| ||||| |||||||||||     ||| ||| | ||  |||||||
Db    3627 ACTTCTTACCAGCGAGAGTAAGCTGAAGCAGCGTGCGCTGACCCTGGAGGTGGAGCG--T 3684

Qy    3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
           |||||||||||||| |||||||||||||||| ||  |||||   |||||||   | ||
Db    3685 CGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGCACTCCAGA 3744

Qy    3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
           ||| || ||||| |||| |||||||| ||||| |||| |||
Db    3745 GCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 3786


RESULT   3
MMU33005
LOCUS       MMU33005                4039 bp    mRNA    linear   ROD 12-JAN-1996
DEFINITION  Mus musculus tbc1 mRNA, complete cds.
ACCESSION   U33005
VERSION     U33005.1  GI:988220
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4039)
  AUTHORS   Richardson,P.M. and Zon,L.I.
  TITLE     Molecular cloning of a cDNA with a novel domain present in the
            tre-2 oncogene and the yeast cell cycle regulators BUB2 and cdc16
  JOURNAL   Oncogene 11 (6), 1139-1148 (1995)
  MEDLINE   96032578
   PUBMED   7566974
REFERENCE   2  (bases 1 to 4039)
  AUTHORS   Richardson,P.M. and Zon,L.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-1995) Leonard I. Zon, Hematology, Childrens
            Hospital of Boston, HHMI, 300 Longwood Ave., Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1. .4039
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_type="mast cells"
     gene            1. .4039
                     /gene="tbc1"
     CDS             403. .3828
                     /gene="tbc1"
                     /note="ORF1"
                     /codon_start=1
                     /product="Tbc1"
                     /protein_id="AAA85223.1"
                     /db_xref="GI:988221"
                     /translation="MPMLPWVVAEVRRLSGQCSKKEPRTKQVRLWVSPSGLRCEPDLE
                     KSQPWDPLICSSIFECKPQRVHKLIHNSHDPSYFACLIKEDAAHRQSLCYVFKADDQT
                     KVPEIISSIRQAGKIARQEELRCPSEFDDTFAKKFEVLFCGRVTVAHKKAPPALIDEC
                     IEKFNHVSCGRRTDWEAPTGQPSAPGPRPMRKSFSQPGLRSLAFRKEFQDASLRSSTF
                     SSFDNDIENHLIGGHNVVQPTDMEENRTMLFTIGPSEVYLISPDTKKIALEKNFKEIS
                     FCSQGIRHVDHFGFICRECSGGGSGGFHFVCYVFQCTNEALVDEIMMTLKQAFTVAAV
                     QQTAKAPAQLCEGCPLQGLHKLCERIEGMNSSKTKLELQKHLTTLTNQEQATIFEEVQ
                     KLRPRNEQRENELIISFLRCLYEEKQKEHSHTGAPKQTLQVAAENIGSDLPPSASRFR
                     LDSLKNRAKRSLTESLESILSRGNKARGLQDHSASVDLDSSTSSTLSNTSKELSMGDK
                     EAFPVSETSFKLLGSSDDLSSDSEGHIAEESALLSPQQAFRRRANTLSHFPVECPAPP
                     EPAQSSPGVSQRKLMRYHSVSTETPHERKDFESKANHLGDTDGTPVKTRRHSWRQQIF
                     LRVATPQKACDSPSRYEDYSELGELPPRSPLEPVCEDGPFGQYRKKRGRRHASFESCG
                     KRPSCSRSCLVRMEKENQKLQASENDLLNKRLKLDYEEITPCLKEVTTVWEKMLSTPG
                     RSKIKFDMEKVHSAVGQGVPRHHRGEIWKFLAEQFHLKHPFPSKQQPKDVPYKELLKK
                     LTSQQHAILIDLGRTFPTHPYFSAQLGAGQLSLYNILKAYSLLDQEVGYCQGLSFVAG
                     ILLLHMSEEEAFKMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHDYHRDLYNHLE
                     EHETGPPTYAAPWFLTVFASQFPLGFVARVFDMIFLQGSEVIFKVALSLLGSHKPLIL
                     QHENLETIVDFIKNTLPNLGLVQMEKTISQVFEMDIAKQLQAYEVEYHVVQEELIESS
                     PLSDNQRMEKLEKTNSTLRKQNLDLLEQLQVANARIQSLEATVEKLLTSESKLKQRAL
                     TLEVERRPAADGGGAAEAKRPAQHSRARLHPAGAHRRLTAARRDCAPLTLSKP"
BASE COUNT     1042 a   1096 c   1060 g    841 t
ORIGIN


  Query Match           70.6%;  Score 2510.4;  DB 10;  Length 4039;
  Best Local Similarity  83.7%;  Pred. No. 0;
  Matches 2949;  Conservative    0;  Mismatches  536;  Indels   37;  Gaps    8;

Qy      1 atggaaccaataacattcacagcaaggaaacatctgctt-cctaacgaggtctcggtgga 59
          |||||  |||| |||||||||||| |||| |||| | || |||||||| ||||| |||||
Db    292 ATGGAGGCAATCACATTCACAGCGAGGAAGCATCCGTTTCCCTAACGAAGTCTCCGTGGA 351

Qy     60 ttttggcctgcagctggt--gggctccctgcctgtgcattccctgaccaccatgcccatg 117
           || ||| |||||||||||  || |||| | || |||||||| || ||||| |||||||||
Db    352 CTTCGGCTTGCAGCTGGTTGGGTCTCCTTACCCGTGCATTCTCTCACCACTATGCCCATG 411

Qy    118 ctgccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacct 177
          |||||  ||||| || || ||||| ||||||||||  | |||| | | | |||||| |||
Db    412 CTGCCGTGGGTGGTAGCCGAGGTACGAAGACTCAGTGGCCAGTGCTCTAAAAAGGAGCCT 471

Qy    178 gtaaccaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagcca 237
             || |||||||||| ||||| ||||| ||||| |||||| | || || ||||| |
Db    472 AGGACAAAGCAAGTCCGCCTTTGGGTTTCGCCCTCCGGACTGCGGTGCGAGCCTGACCTG 531

Qy    238 gggagaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcag 297
          | || |||||||  ||||| || || |||| ||||||||||||||||||||||||||||||
Db    532 GAGAAAAGTCAACCATGGGACCCGCTCATCTGTTCCAGCATCTTTGAGTGCAAGCCTCAG 591

Qy    298 cgtgttcacaaactgattcacaacagtcatgacccaagttactttgcttgtctgattaag 357
          |||||  |||||||||||||||||||||||| ||||| || |||||||||| || ||||||
Db    592 CGTGTGCACAAACTGATTCACAACAGTCACGACCCGAGCTACTTTGCTTGCCTCATTAAG 651

Qy    358 gaagacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaa 417
          || || ||||  ||| |||||||  |||||||||||||| ||||| || ||||||||||||
Db    652 GAGGATGCTGCGCACAGGCAGAGCCTCTGCTATGTGTTTAAAGCAGACGATCAAACAAAA 711

Qy    418 gtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagctg 477
          ||||||||||||||||||||||||||||| ||||| |||||||| ||||| ||||| ||||||
Db    712 GTGCCTGAGATCATCAGCTCCATCCGGCAGGCCGGGAAGATTGCCCGCCAGGAAGAGCTG 771

Qy    478 cactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcggc 537
          |  ||||| |||||||||||||| || ||  |||| ||||||||||||||||||||| |||
Db    772 CGTTGCCCCTCCGAGTTCGACGATACCTTCGCCAAAAAGTTCGAGGTGCTCTTCTGTGGC 831

Qy    538 cgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgagaag 597
          || ||||| ||||| |||||||||| || || || ||||| |||||||| |||||||||
Db    832 CGGGTGACTGTGGCTCACAAGAAGGCCCCACCCGCACTGATTGACGAGTGTATCGAGAAG 891

Qy    598 ttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatgcc 657
          ||||| || || ||| |  | ||  |  | ||  |       |||| |  ||  |   |
```



```
Qy   2811 gtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggagga 2870
          ||||||||| ||  || | || || |||||||| |||||||||||| |||||||| ||
Db   3087 GTACCAGCTGTCACGGCTCCTCCACGATTACCACCGAGACCTCTACAACCACCTGGAAGA 3146

Qy   2871 gcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcaca 2930
          ||||||||  ||||||     ||||| || |||||||| ||||||  |||| ||||||||
Db   3147 GCACGAGACTGGCCCCCCTACGTACGCGGCTCCCTGGTTTCTCACCGTGTTCGCCTCACA 3206

Qy   2931 gttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggt 2990
          |||||| || || || ||||||||||||||||||||||| || |||||||| |||||||
Db   3207 GTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGATCAGAGGT 3266

Qy   2991 catatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatga 3050
          |||||||||||| |||||||||| ||||| |||||||||||||||||||| ||||||||
Db   3267 CATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTACAGCATGA 3326

Qy   3051 aaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtaca 3110
           ||||| |||||||| || ||||| ||||| | ||| || |||||||| ||| |||| ||
Db   3327 GAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGCCTGGTGCA 3386

Qy   3111 gatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagctta 3170
          |||||| |||||||||| |||||| |||| |||||||||||| || ||| | || || ||
Db   3387 GATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTCCAGGCCTA 3446

Qy   3171 tgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagtgacaa 3230
          ||| || |||||||||||| | ||| || || ||||| || ||||| ||||||||||||||
Db   3447 TGAGGTCGAGTACCACGTCGTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTCAGTGACAA 3506

Qy   3231 ccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgacct 3290
          |||||||||||| ||||| ||||||||||||||||  || ||||||||||||||||||||
Db   3507 CCAAAGAATGGAGAAATTGGAGAAAACCAACAGCACGTTGCGCAAACAGAACCTTGACCT 3566

Qy   3291 ccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa 3350
          ||| || ||||||||||||||||||| |||||||||||||||||| |||||  | |||||
Db   3567 CCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACGGTAGAGAA 3626

Qy   3351 gctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
          || || | ||| ||||| ||||||||||| .  ||| ||| | ||  |||||||
Db   3627 ACTTCTTACCAGCGAGAGTAAGCTGAAGCAGCGTGCGCTGACCCTGGAGGTGGAGCG--T 3684

Qy   3411 ggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcgaccggga 3470
          |||||||||||| |||||||||||||||| ||| |||||   |||||||   |  ||
Db   3685 CGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGCACTCCAGA 3744

Qy   3471 gcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512
          ||| || ||||| |||| |||||||| ||||| |||| |||
Db   3745 GCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 3786


RESULT   4
AB029031
LOCUS       AB029031               2576 bp    mRNA    linear   PRI 04-AUG-1999
DEFINITION  Homo sapiens mRNA for KIAA1108 protein, partial cds.
ACCESSION   AB029031
VERSION     AB029031.1  GI:5689552
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hh03387.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kikuno,R., Nagase,T., Ishikawa,K., Hirosawa,M., Miyajima,N.,
            Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XIV. The complete sequences of 100 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 6 (3), 197-205 (1999)
  MEDLINE   99397452
REFERENCE   2  (bases 1 to 2576)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1. .2576
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /clone="hh03387"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1. .2292
                     /gene="KIAA1108"
     CDS             <1. .2292
                     /gene="KIAA1108"
                     /note="hh03387 cDNA clone for KIAA1108 has a 3306-bp
                     insertion."
                     /codon_start=1
                     /product="KIAA1108 protein"
                     /protein_id="BAA83060.1"
                     /db_xref="GI:5689553"
                     /translation="NSSKTKLELQKHLTTLTNQEQATIFEEVQKLRPRNEQRENELII
                     SFLRCLYEEKQKEHIHIGEMKQTSQMAAENIGSELPPSATRFRLDMLKNKAKRSLTES
                     LESILSRGNKARGLQEHSISVDLDSSLSSTLSNTSKEPSVCEKEALPISESSFKLLGS
                     SEDLSSDSESHLPEEPAPLSPQQAFRRRANTLSHFPIECQEPPQPARGSPGVSQRKLM
                     RYHSVSTETPHERKDFESKANHLGDSGGTPVKTRRHSWRQQIFLRVATPQKACDSSSR
                     YEDYSELGELPPRSPLEPVCEDGPFGPPPEEKKRTSRELRELWQKAILQQILLLRMEK
                     ENQKLQASENDLLNKRLKLDYEEITPCLKEVTTVWEKMLSTPGRSKIKFDMEKMHSAV
                     GQGVPRHHRGEIWKFLAEQFHLKHQFPSKQQPKDVPYKELLKQLTSQQHAILIDLGRT
                     FPTHPYFSAQLGAGQLSLYNILKAYSLLDQEVGYCQGLSFVAGILLLHMSEEEAFKML
                     KFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHDYHRDLYNHLEEHEIGPSLYAAPWFL
                     TMFASQFPLGFVARVFDMIFLQGTEVIFKVALSLLGSHKPLILQHENLETIVDFIKST
                     LPNLGLVQMEKTINQVFEMDIAKQLQAYEVEYHVLQEELIDSSPLSDNQRMDKLEKTN
                     SSLRKQNLDLLEQLQVANGRIQSLEATIEKLLSSESKLKQAMLTLELERSALLQTVEE
                     LRRRSAEPSDREPECTQPEPTGD"
BASE COUNT      746 a    615 c    628 g    587 t
ORIGIN


  Query Match             65.8%;  Score 2340;  DB 9;  Length 2576;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2340;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1216 aattcttccaaaacaaaactagaactgcaaaagcacctgacgacattaaccaatcaggag 1275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AATTCTTCCAAAACAAAACTAGAACTGCAAAAGCACCTGACGACATTAACCAATCAGGAG 60

Qy   1276 caggcgactatttttgaagaggttcagaaattgagaccgagaaatgagcagcgagagaat 1335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGCGACTATTTTTGAAGAGGTTCAGAAATTGAGACCGAGAAATGAGCAGCGAGAGAAT 120

Qy   1336 gaattgattatttctttttctgagatgtttatatgaagagaaacagaaagaacacatccat 1395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAATTGATTATTTCTTTTCTGAGATGTTTATATGAAGAGAAACAGAAAGAACACATCCAT 180

Qy   1396 attggggagatgaagcagacatcgcagatggcagcagagaatattggaagtgaattacca 1455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATTGGGGAGATGAAGCAGACATCGCAGATGGCAGCAGAGAATATTGGAAGTGAATTACCA 240

Qy   1456 cccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaagagatctttaaca 1515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCAGTGCCACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAGAGATCTTTAACA 300

Qy   1516 gagtctttagaaagtattttgtcccggggtaataaagccagaggcctgcaggaacactcc 1575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGTCTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTGCAGGAACACTCC 360

Qy   1576 atcagtgtggatctggatagctccctgtctagtacattaagtaacaccagcaaagagcca 1635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCAGTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACCAGCAAAGAGCCA 420

Qy   1636 tctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaagctcctcggctcc 1695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    421 TCTGTGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCC 480

Qy   1696 tcggaggacctgtccagtgactcggagagtcatctcccagaagagccagctccgctgtcg 1755
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCAGCTCCGCTGTCG 540

Qy   1756 ccccagcaggccttcaggaggcgagcaaacaccctgagtcacttccccatcgaatgccag 1815
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCCATCGAATGCCAG 600

Qy   1816 gaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaacttatgaggtat 1875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAACTTATGAGGTAT 660

Qy   1876 cactcagtgagcacagagacgcctcatgaacgaaaggactttgaatccaaagcaaaccat 1935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCCAAAGCAAACCAT 720

Qy   1936 cttggtgattctggtgggactcctgtgaagacccggaggcattcctggaggcagcagata 1995
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGGAGGCAGCAGATA 780

Qy   1996 ttcctccgagtagccaccccgcagaaggcgtgcgattcttccagcagatatgaagattat 2055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGATATGAAGATTAT 840

Qy   2056 tcagagctgggagagcttcccccacgatctcctttagaaccagtttgtgaagatgggccc 2115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCC 900

Qy   2116 tttggcccccccaccagaggaaaagaaaaggacatctcgtgagctccgagagctgtggcaa 2175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAA 960

Qy   2176 aaggctattcttcaacagatactgctgcttagaatggagaaggaaaatcagaagctccaa 2235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAATCAGAAGCTCCAA 1020

Qy   2236 gcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaagaaattactccc 2295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAAATTACTCCC 1080

Qy   2296 tgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactccaggaagatcaaaa 2355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCAGGAAGATCAAAA 1140

Qy   2356 attaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtgccacgtcatcac 2415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTGCCACGTCATCAC 1200

Qy   2416 cgaggtgaaatctggaaatttctagctgagcaattccaccttaaacaccagtttcccagc 2475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACACCAGTTTCCCAGC 1260

Qy   2476 aaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctgacttcccagcag 2535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAG 1320

Qy   2536 catgcgattcttattgaccttgggcgaaccttcctacacacccatacttctctgcccag 2595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAG 1380

Qy   2596 cttggagcaggacagctatcgctttacaacattttgaaggcctactcacttctagaccag 2655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCACTTCTAGACCAG 1440

Qy   2656 gaagtgggatattgccaaggtctcagctttgtagcaggcattttgcttcttcatatgagt 2715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTTTGCTTCTTCATATGAGT 1500

Qy   2716 gaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatggggctgcggaaacag 2775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAAACAG 1560
```

```
Qy   2776 tatcggccagacatgattattttacagatccagatgtaccagctctcgaggttgcttcat 2835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCTTCAT 1620

Qy   2836 gattaccacagagacctctacaatcacctggaggagcacgagatcggccccagcctctac 2895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCTCTAC 1680

Qy   2896 gctgccccctggttcctcaccatgtttgcctcacagttcccgctgggattcgtagccaga 2955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCAGA 1740

Qy   2956 gtctttgatatgatttttcttcagggaacagaggtcatatttaaagtggctttaagtctg 3015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTCTG 1800

Qy   3016 ttgggaagccataagcccttgattctgcagcatgaaaacctagaaaccatagttgacttt 3075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACTTT 1860

Qy   3076 ataaaaagcacgctacccaaccttggcttggtacagatggaaaagaccatcaatcaggta 3135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGGTA 1920

Qy   3136 tttgaaatggacatcgctaaacagttacaagcttatgaagttgagtaccacgtccttcaa 3195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TTTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTCAA 1980

Qy   3196 gaagaacttatcgattcctctcctctcagtgacaaccaaagaatggataaattagagaaa 3255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGAAA 2040

Qy   3256 accaacagcagcttacgcaaacagaaccttgacctccttgaacagttgcaggtggcaaat 3315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCAACAGCAGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAAAT 2100

Qy   3316 ggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagctg 3375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GGTAGGATCCAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGCTG 2160

Qy   3376 aagcaggccatgcttaccttagaactggagcggtcggccctgctgcagacggtggaggag 3435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AAGCAGGCCATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGGAG 2220

Qy   3436 ctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacgcagcccgagccc 3495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CTGCGGCGGCGGAGCGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGCCC 2280

Qy   3496 acgggcgactgacagctctgcaggagagattgcaacaccatcccacactgtccaggcctt 3555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 ACGGGCGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCCTT 2340
```

```
RESULT   5
AK027355
LOCUS       AK027355               2362 bp    mRNA    linear   PRI 15-MAY-2001
DEFINITION  Homo sapiens cDNA FLJ14449 fis, clone HEMBB1001684.
ACCESSION   AK027355
VERSION     AK027355.1  GI:14041975
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens embryo, 10 weeks whole embryo, mainly body cDNA to
            mRNA, clone_lib:HEMBB1 clone:HEMBB1001684.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
            Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
```

```
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2362)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology; cDNA library construction,
            5'- & 3'-end one pass sequencing and clone selection: Helix
            Research Institute (supported by Japan Key Technology Center etc.)
            and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .2362
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBB1001684"
                     /tissue_type="whole embryo, mainly body"
                     /clone_lib="HEMBB1"
                     /dev_stage="embryo, 10 weeks"
                     /note="cloning vector: pME18SFL3"
     CDS             40. .2064
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB55057.1"
                     /db_xref="GI:14041976"
                     /translation="MLKNKAKRSLTESLESILSRGNKARGLQEHSISVDLDSSLSSTL
                     SNTSKEPSVCEKEALPISESSFKLLGSSEDLSSDSESHLPEEPAPLSPQQAFRRRANT
                     LSHFPIECQEPPQPARGSPGVSQRKLMRYHSVSTETPHERKDFESKANHLGDSGGTPV
                     KTRRHSWRQQIFLRVATPQKACDSSSRYEDYSELGELPPRSPLEPVCEDGPFGPPPEE
                     KKRTSRELRELWQKAILQQILLLRMEKENQKLQASENDLLNKRLKLDYEEITPCLKEV
                     TTVWEKMLSTPGRSKIKFDMEKMHSAVGQGVPRHHRGEIWKFLAEQFHLKHQFPSKQQ
                     PKDVPYKELLKQLTSQQHAILIDLGRTFPTHPYFSAQLGAGQLSLYNILKAYSLLDQE
                     VGYCQGLSFVAGILLLHMSEEEAFKMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLL
                     HDYHRDLYNHLEEHEIGPSLYAAPWFLTMFASQFPLGFVARVFDMIFLQGTEVIFKVA
                     LSLLGSHKPLILQHENLETIVDFIKSTLPNLGLVQMEKTINQVFEMDIAKQLQAYEVE
                     YHVLQEELIDSSPLSDNQRMDKLEKTNSSLRKQNLDLLEQLQVANGRIQSLEATIEKL
                     LSSESKLKQAMLTLELERSALLQTVEELRRRSAEPSDREPECTQPEPTGD"
BASE COUNT      657 a    581 c    578 g    546 t
ORIGIN


  Query Match           59.4%;  Score 2112;  DB 9;  Length 2362;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2112;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1444 agtgaattaccacccagtgccactcgatttaggctagatatgctgaaaaacaaagcaaag 1503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGTGAATTACCACCCAGTGCCACTCGATTTAGGCTAGATATGCTGAAAAACAAAGCAAAG 60

Qy   1504 agatctttaacagagtctttagaaagtattttgtcccggggtaataaagccagaggcctg 1563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGATCTTTAACAGAGTCTTTAGAAAGTATTTTGTCCCGGGGTAATAAAGCCAGAGGCCTG 120

Qy   1564 caggaacactccatcagtgtggatctggatagctccctgtctagtacattaagtaacacc 1623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGGAACACTCCATCAGTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAGTAACACC 180

Qy   1624 agcaaagagccatctgtgtgtgaaaaggaggccttgcccatctctgagagctcctttaag 1683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCAAAGAGCCATCTGTGTGTGAAAAGGAGGCCTTGCCCATCTCTGAGAGCTCCTTTAAG 240

Qy   1684 ctcctcggctcctcggaggacctgtccagtgactcggagagtcatctcccagaagagcca 1743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTCCTCGGCTCCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAGAAGAGCCA 300

Qy   1744 gctccgctgtcgccccagcaggccttcaggaggcgagcaaacaccctgagtcacttcccc 1803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCTCCGCTGTCGCCCCAGCAGGCCTTCAGGAGGCGAGCAAACACCCTGAGTCACTTCCCC 360

Qy   1804 atcgaatgccaggaacctccacaacctgcccgggggtccccgggggtttcgcaaaggaaa 1863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCGAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTTCGCAAAGGAAA 420

Qy   1864 cttatgaggtatcactcagtgagcacagagacgcctcatgaacgaaaggactttgaatcc 1923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTTATGAGGTATCACTCAGTGAGCACAGAGACGCCTCATGAACGAAAGGACTTTGAATCC 480

Qy   1924 aaagcaaaccatcttggtgattctggtgggactcctgtgaagacccggaggcattcctgg 1983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAAGCAAACCATCTTGGTGATTCTGGTGGGACTCCTGTGAAGACCCGGAGGCATTCCTGG 540

Qy   1984 aggcagcagatattcctccgagtagccaccccgcagaaggcgtgcgattcttccagcaga 2043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGGCAGCAGATATTCCTCCGAGTAGCCACCCCGCAGAAGGCGTGCGATTCTTCCAGCAGA 600

Qy   2044 tatgaagattattcagagctgggagagcttcccccacgatctcctttagaaccagtttgt 2103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCCACGATCTCCTTTAGAACCAGTTTGT 660

Qy   2104 gaagatgggccctttggccccccaccagaggaaaagaaaaggacatctcgtgagctccga 2163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAAGATGGGCCCTTTGGCCCCCCACCAGAGGAAAAGAAAAGGACATCTCGTGAGCTCCGA 720

Qy   2164 gagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggagaaggaaaat 2223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGCTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAGAAGGAAAAT 780

Qy   2224 cagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctcgattatgaa 2283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAGAAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAA 840

Qy   2284 gaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgcttagcactcca 2343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GAAATTACTCCCTGTCTTAAAGAAGTAACTACAGTGTGGGAAAAGATGCTTAGCACTCCA 900

Qy   2344 ggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgttgggcaaggtgtg 2403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGAAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTTGGGCAAGGTGTG 960

Qy   2404 ccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattccaccttaaacac 2463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCACGTCATCACCGAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCCACCTTAAACAC 1020

Qy   2464 cagtttcccagcaaacagcagccaaaggatgtgccatacaaagaactcttaaagcagctg 2523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAGTTTCCCAGCAAACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTG 1080

Qy   2524 acttcccagcagcatgcgattcttattgaccttgggcgaacctttcctacacacccatac 2583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACTTCCCAGCAGCATGCGATTCTTATTGACCTTGGGCGAACCTTTCCTACACACCCATAC 1140

Qy   2584 ttctctgcccagcttggagcaggacagctatcgctttacaacattttgaaggcctactca 2643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTCTCTGCCCAGCTTGGAGCAGGACAGCTATCGCTTTACAACATTTTGAAGGCCTACTCA 1200

Qy   2644 cttctagaccaggaagtgggatattgccaaggtctcagctttgtagcaggcattttgctt 2703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTTTGCTT 1260

Qy   2704 cttcatatgagtgaggaagaggcgtttaaaatgctcaagtttctgatgtttgacatgggg 2763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTTCATATGAGTGAGGAAGAGGCGTTTAAAATGCTCAAGTTTCTGATGTTTGACATGGGG 1320

Qy   2764 ctgcggaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 2823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGCGGAAACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCG 1380

Qy   2824 aggttgcttcatgattaccacagagacctctacaatcacctggaggagcacgagatcggc 2883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AGGTTGCTTCATGATTACCACAGAGACCTCTACAATCACCTGGAGGAGCACGAGATCGGC 1440
```

```
Qy   2884 cccagcctctacgctgccccctggttcctcaccatgtttgcctcacagttcccgctggga 2943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCCAGCCTCTACGCTGCCCCCTGGTTCCTCACCATGTTTGCCTCACAGTTCCCGCTGGGA 1500

Qy   2944 ttcgtagccagagtctttgatatgatttttcttcagggaacagaggtcatatttaaagtg 3003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTCGTAGCCAGAGTCTTTGATATGATTTTTCTTCAGGGAACAGAGGTCATATTTAAAGTG 1560

Qy   3004 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaacctagaaacc 3063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCTTTAAGTCTGTTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACC 1620

Qy   3064 atagttgacttttataaaaagcacgctacccaaccttggcttggtacagatggaaaagacc 3123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ATAGTTGACTTTATAAAAAGCACGCTACCCAACCTTGGCTTGGTACAGATGGAAAAGACC 1680

Qy   3124 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttatgaagttgagtac 3183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATCAATCAGGTATTTGAAATGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTAC 1740

Qy   3184 cacgtccttcaagaagaacttatcgattcctctcctctcagtgacaaccaaagaatggat 3243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CACGTCCTTCAAGAAGAACTTATCGATTCCTCTCCTCTCAGTGACAACCAAAGAATGGAT 1800

Qy   3244 aaattagagaaaaccaacagcagcttacgcaaacagaaccttgacctccttgaacagttg 3303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AAATTAGAGAAAACCAACAGCAGCTTACGCAAACAGAACCTTGACCTCCTTGAACAGTTG 1860

Qy   3304 caggtggcaaatggtaggatccaaagccttgaggccaccattgagaagctcctgagcagt 3363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 CAGGTGGCAAATGGTAGGATCCAAAGCCTTGAGGCCACCATTGAGAAGCTCCTGAGCAGT 1920

Qy   3364 gagagcaagctgaagcaggccatgcttaccttagaactggagcggtcggccctgctgcag 3423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GAGAGCAAGCTGAAGCAGGCCATGCTTACCTTAGAACTGGAGCGGTCGGCCCTGCTGCAG 1980

Qy   3424 acggtggaggagctgcggcggcggagcgcagagcccagcgaccgggagcctgagtgcacg 3483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACGGTGGAGGAGCTGCGGCGGCGGAGCGCAGAGCCCAGCGACCGGGAGCCTGAGTGCACG 2040

Qy   3484 cagcccgagcccacgggcgactgacagctctgcaggagagattgcaacaccatcccacac 3543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 CAGCCCGAGCCCACGGGCGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCCACAC 2100

Qy   3544 tgtccaggcctt 3555
          ||||||||||||
Db   2101 TGTCCAGGCCTT 2112
```

```
RESULT   6
BTA17923
LOCUS       BTA17923                1995 bp    mRNA    linear   MAM 18-NOV-1998
DEFINITION  Bos taurus mRNA for lyncein.
ACCESSION   Y17923
VERSION     Y17923.1  GI:3893076
KEYWORDS    lyn gene; lyncein.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 1995)
  AUTHORS   Xu,X.Z., Wes,P.D., Chen,H., Li,H.S., Yu,M., Morgan,S., Liu,Y. and
            Montell,C.
  TITLE     Retinal targets for calmodulin include proteins implicated in
            synaptic transmission
  JOURNAL   J. Biol. Chem. 273 (47), 31297-31307 (1998)
  MEDLINE   99030403
REFERENCE   2  (bases 1 to 1995)
  AUTHORS   Montell,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-1998) C. Montell, The Johns Hopkins University
            School of Medicine, Department of Biological Chermistry, 725 N.
            Wolfe Street, Baltimore, MD 21205-2185, USA
FEATURES             Location/Qualifiers
     source          1. .1995
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="retinal"
                     /dev_stage="adult"
     gene            2. .1672
                     /gene="lyn"
     CDS             <2. .1672
                     /gene="lyn"
                     /codon_start=1
                     /product="lyncein"
                     /protein_id="CAA76943.1"
                     /db_xref="GI:3893077"
                     /translation="ARAGVSQRKLVRYHSVSTETPHERKDFESKADHISDASRTPVKT
                     RRHSWRQQIFLRVATPQKACESPKRYEDYSELGELPPRSPLEPVCEDGPFGPVPEEKK
                     RTSHELRELWQKAILQQILLLRMEKENQKLQASENDLLNKRLKLDYEEITPCLKEVTT
                     VWEKILSTPGRSKIKFDMEKMHSAVGQGVPRHHRGEIWKFLAEQYHLKHPFPCKQQPK
                     DTPYKELLKQLTSQQHAILIDLGRTFPTHPYYSAQLGAGQLSLYNILKAYSLLDQEVG
                     YCQGLSFVAGILLLHMGEEEAFNMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHD
                     YHRDLYNHLEEHEIGPSLYAAPWFLTVFASQFPLGFVARVFDMIFLQGSEVIFKVALS
                     LLGSHKPLILQHENLETIVDFIKSTLPNLGLVQMEKTISQVFETDISKQLQAYEVEYH
                     VLQEELIDSSPLSDNQRMDKLEKTNSSLRKQNLDLLEQLQVANGRIQSLEATVEKLLT
                     SESKLKQATLALELERSALLQTVEQLRRQTAELGSQESDPTLPKPSGD"
BASE COUNT      560 a    516 c    497 g    422 t
ORIGIN
```

Query Match 38.2%; Score 1359.2; DB 4; Length 1995;
Best Local Similarity 88.4%; Pred. No. 1.9e-286;
Matches 1475; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

```
Qy   1844 cgggggtttcgcaaaggaaacttatgaggtatcactcagtgagcacagagacgcctcatg 1903
          | |||||  ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      9 CAGGGGTCTCGCAAAGGAAACTTGTGAGGTATCACTCAGTGAGCACAGAGACGCCTCATG 68

Qy   1904 aacgaaaggactttgaatccaaagcaaaccatcttggtgattctggtgggactcctgtga 1963
          ||||||||||||||| ||||||||  ||||  |  | |||  |  |  ||||  |||||||
Db     69 AACGAAAGGACTTTGAGTCCAAAGCCGACCACATCAGCGATGCCAGCAGGACCCCTGTGA 128

Qy   1964 agacccggaggcattcctggaggcagcagatattcctccgagtagccaccccgcagaaagg 2023
          ||||  |||||  ||  ||  ||||||||||||||  |||||  ||  ||  ||||||||  |||||||
Db    129 AGACACGGAGACACTCGTGGAGGCAGCAGATCTTCCTTCGGGTGGCCACCCCCCAGAAGG 188

Qy   2024 cgtgcgattcttccagcagatatgaagattattcagagctgggagagcttcccccacgat 2083
          | ||||| ||  |||  |||||  ||||||||  ||||||  ||||||||||||||||||||||
Db    189 CCTGCGAGTCCCCCAAAAGATACGAAGATTACTCAGAGTTGGGAGAGCTTCCCCCACGAT 248

Qy   2084 ctcctttagaaccagtttgtgaagatgggccctttggccccccaccagaggaaaagaaaa 2143
          ||||  ||||||||||||||||||||  ||||||||||    ||||||||||||||||||
Db    249 CTCCACTAGAACCAGTTTGTGAAGATGGACCCTTTGGCCCGGTCCCAGAGGAAAAGAAAA 308

Qy   2144 ggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctgc 2203
          | |||||||  |||  ||  ||  ||||||||||||||||||  |||||  |||||||||  ||||||
Db    309 GAACATCTCATGAACTTCGGGAGCTGTGGCAAAAGGCCATTCTACAACAGATATTGCTGC 368

Qy   2204 ttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagc 2263
          | ||  ||||||||  |||||  ||  ||  ||  ||||||||||  ||||||||||||||||||
Db    369 TCAGGATGGAGAAAGAAAACCAAAAACTGCAAGCCTCTGAGAATGATTTGCTGAACAAGC 428

Qy   2264 gcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtggg 2323
          ||||  |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    429 GCCTAAAGCTTGATTATGAAGAAATCACTCCCTGTCTTAAAGAAGTGACTACAGTGTGGG 488

Qy   2324 aaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcact 2383
          |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    489 AAAAGATACTTAGCACTCCAGGAAGATCAAAAATTAAGTTTGACATGGAAAAAATGCACT 548

Qy   2384 cggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctg 2443
          ||||||||||  ||||||||||  ||  ||  |||||  ||||||||||||||||  |  ||  |
Db    549 CGGCTGTTGGCCAAGGTGTGCCGCGACACCACCGGGGTGAAATCTGGAAATTCTTGGCGG 608
```

```
Qy   2444 agcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccataca 2503
          |||| | ||||||  || ||||| |||||||  ||||||||||||| |||||   ||| ||||
Db    609 AGCAGTACCACCTGAAGCACCCGTTTCCCTGCAAACAGCAGCCCAAGGACACGCCCTACA 668

Qy   2504 aagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaa 2563
          ||||  |||||||||||||| || |||||||||| || ||||| || ||||||||||||
Db    669 AAGAGCTCTTAAAGCAGCTCACCTCCCAGCAGCACGCCATTCTCATCGACCTTGGGCGAA 728

Qy   2564 cctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttaca 2623
          ||||||| ||||| || |||| |||||  ||||| |||||||| ||||| || |||||||
Db    729 CCTTTCCAACACATCCGTACTACTCTGCTCAGCTAGGAGCAGGGCAGCTGTCACTTTACA 788

Qy   2624 acattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagct 2683
          ||||  |||||||||||| |||||||||| |||||||| ||||||||||||||||||||
Db    789 ACATCCTGAAGGCCTACTCGCTTCTAGACCAAGAAGTGGGCTATTGCCAAGGTCTCAGCT 848

Qy   2684 ttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagt 2743
          ||||||||||||  |||| |||||||| |||| |||||||||||||| ||||||||||
Db    849 TTGTAGCAGGCATCCTGCTGCTTCATATGGGTGAAGAAGAGGCGTTTAACATGCTCAAGT 908

Qy   2744 ttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacaga 2803
          |||||||||||||||||| |||||||||| |||||||| ||||||||||||||||||||
Db    909 TTCTGATGTTTGACATGGGACTGCGGAAACAATATCGGCCGGACATGATTATTTTACAGA 968

Qy   2804 tccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacc 2863
          ||||||||||||||||||||||| |||| || || |||||||| |||||||||| ||||
Db    969 TCCAGATGTACCAGCTCTCGAGGCTGCTCCACGACTACCACAGGGACCTCTACAACCACC 1028

Qy   2864 tggaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttg 2923
          |||||||||| ||||| |||||||||||||| ||||| |||||||||||||| ||||||
Db   1029 TGGAGGAGCATGAGATTGGCCCCAGCCTCTATGCTGCGCCCTGGTTCCTCACCGTGTTTG 1088

Qy   2924 cctcacagttcccgctcggattcgtagccagagtctttgatatgatttttcttcagggaa 2983
          |||||||||||| ||||| ||||| |||||||||||||||||||||| ||||||||||||
Db   1089 CCTCACAGTTCCCCCTGGGCTTCGTCGCCAGAGTCTTTGATATGATCTTTCTTCAGGGAT 1148

Qy   2984 cagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctgc 3043
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1149 CAGAGGTCATATTTAAAGTGGCTTTAAGTCTGCTGGGAAGCCATAAGCCCTTGATTCTGC 1208

Qy   3044 agcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggct 3103
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||
Db   1209 AGCATGAAAACCTAGAAACCATAGTTGACTTTATAAAAAGCACACTACCCAACCTGGGCT 1268

Qy   3104 tggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttac 3163
          |||| |||||||| |||| |||| |||||  ||||| | ||||||| | ||||||||||
Db   1269 TGGTGCAGATGGAGAAGACTATCAGTCAGGTGTTTGAGACGGACATCTCGAAACAGTTAC 1328

Qy   3164 aagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctca 3223
          |||||||||||||||||||||| ||||| |||||  |||||||||| ||||||||||
Db   1329 AAGCTTATGAAGTTGAGTACCACGTGCTTCAGGAAGAGCTTATCGATTCTTCTCCTCTCA 1388

Qy   3224 gtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaacc 3283
          |||||||||||||||||| ||| ||||||| ||||||||||||||||||||||||||||
Db   1389 GTGACAACCAAAGAATGGACAAACTAGAGAAGACCAACAGCAGCTTACGCAAACAGAACC 1448

Qy   3284 ttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccacca 3343
          |||||||||||||||||| |||||||| || ||||| |||||||| ||||| || |||||
Db   1449 TTGACCTCCTTGAACAGCTGCAGGTCGCCAATGGGAGGATCCAGAGCCTGGAAGCCACTG 1508

Qy   3344 ttgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactgg 3403
          ||||||||||||||  || |||||||||||||||||| |  ||  |||| || ||||
Db   1509 TTGAGAAGCTCCTGACTAGCGAGAGCAAGCTGAAGCAGGCGACCCTGGCCTTGGAGCTGG 1568

Qy   3404 agcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagcg 3463
          || |||| |||||||||||||||||||||| |||| ||| ||| ||  |||||||   ||
Db   1569 AGAGGTCAGCCCTGCTGCAGACGGTGGAGCAGCTCCGGAGGCAGACGGCAGAGCTGGGCA 1628

Qy   3464 accgggagcctgagtgcacgcagcccgagcccacgggcgactgacagc 3511
            || |||| ||||   ||| | |||| ||||| | || ||||||  ||
Db   1629 GCCAGGAGTCTGACCCCACACTGCCCAAGCCCTCAGGTGACTGATGGC 1676
```

```
RESULT   7
BC004675
LOCUS       BC004675                3051 bp    mRNA    linear   ROD 19-JUL-2001
DEFINITION  Mus musculus, clone IMAGE:3500261, mRNA, partial cds.
ACCESSION   BC004675
VERSION     BC004675.1  GI:14919386
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3051)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gibbs,R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 9 Row: m Column: 4.
FEATURES             Location/Qualifiers
     source          1. .3051
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3500261"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             <1. .1485
                     /codon_start=1
                     /product="Unknown (protein for IMAGE:3500261)"
                     /protein_id="AAH04675.1"
                     /db_xref="GI:14919387"
                     /translation="ACDSPSRYEDYSELGELPPRSPLEPVCEDGPFGPVQEEKRKTSR
                     ELRELWKKAILQQILLLRMEKENQKLQASENDLLNKRLKLDYDEITPCLKEVTTVWEK
                     MLSTPGRSKIKFDMEKVHSAVGQGVPRHHRGEIWKFLAEQFHLKHPFPSKQQPKDVPY
                     KELLKKLTSQQHAILIDLGRTFPTHPYFSAQLGAGQLSLYNILKAYSLLDQEVGYCQG
                     LSFVAGILLLHMSEEEAFKMLKFLMFDMGLRKQYRPDMIILQIQMYQLSRLLHDYHRD
                     LYNHLEEHEIGPSLYAAPWFLTVFASQFPLGFVARVFDMIFLQGSEVIFKVALSLLGS
                     HKPLILQHENLETIVDFIKNTLPNLGLVQMEKTISQVFEMDIAKQLQPYEVEYHVLQE
                     ELIESSPLSDNQRMEKLEKTNSSLRKQNLDLLEQLQVANARIQSLEATVEKLLTSESK
                     LKQAALTLEVERSALLQMVEELRRQSARPSTPEPDCTQLEPTGD"
BASE COUNT      817 a    715 c    756 g    763 t
ORIGIN
```

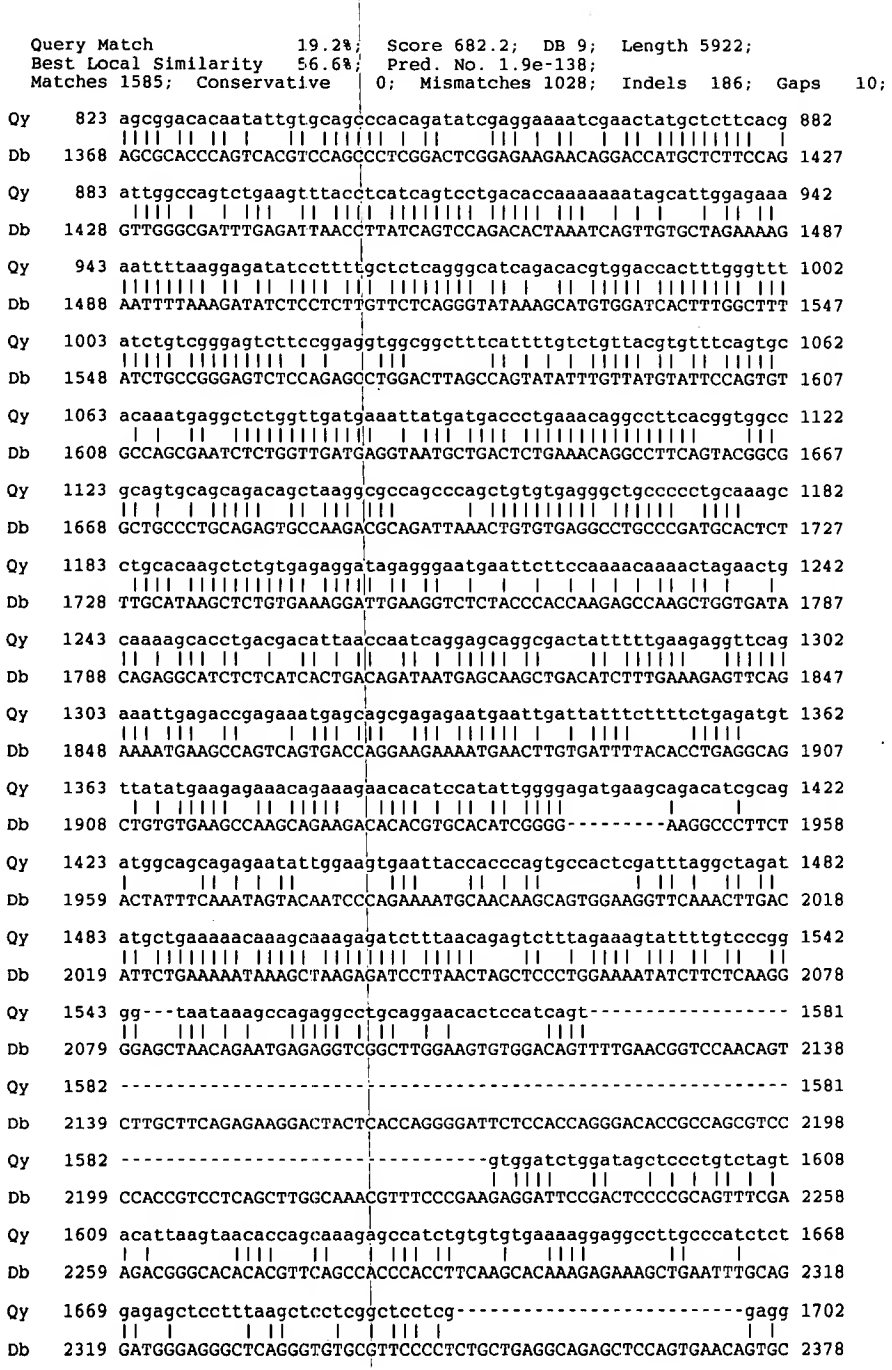
```
  Query Match             33.2%;  Score 1179.6;  DB 10;  Length 3051;
  Best Local Similarity   87.0%;  Pred. No. 3.5e-247;
  Matches 1296;  Conservative    0;  Mismatches  194;  Indels    0;  Gaps    0;

Qy   2023 gcgtgcgattcttccagcagatatgaagattattcagagctgggagagcttcccccacga 2082
          || || || ||  | ||||||||||||||||||| |||||||||||||| || |||||
Db      1 GCTTGTGACTCCCCGAGCAGATATGAAGATTATTCCGAGCTGGGAGAGCTCCCTCCACGC 60

Qy   2083 tctcctttagaaccagtttgtgaagatgggccctttggcccccccaccagaggaaaagaaa 2142
```

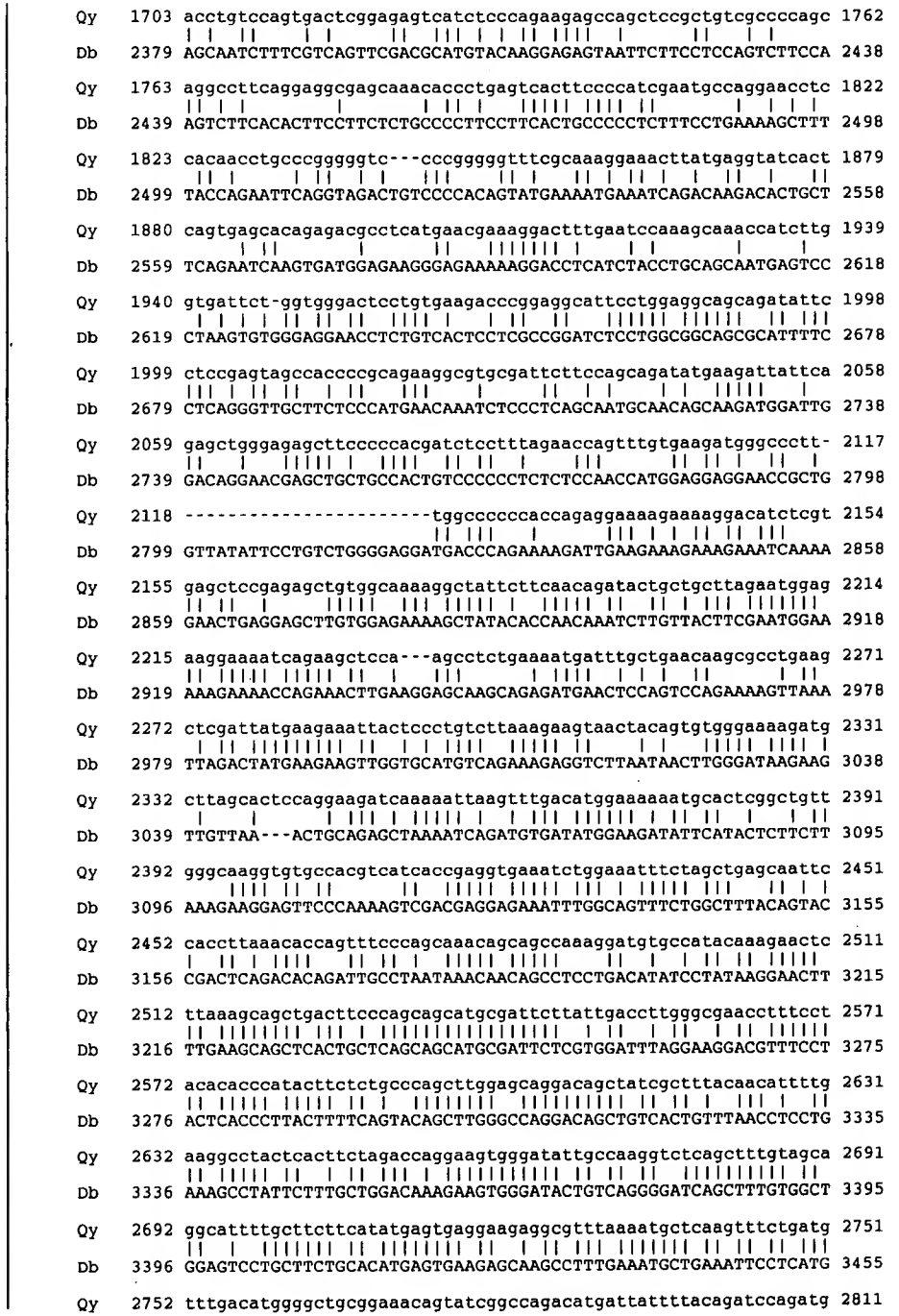
```
Db      61 TCCCCTTTAGAACCGGTGTGTGAGGACGGCCCATTTGGCCCAGTACAGGAAGAAAAGAGG 120

Qy    2143 aggacatctcgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgctg 2202

Db     121 AAGACGTCACGCGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTGCAGCAGATCCTGCTG 180

Qy    2203 cttagaatggagaaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaag 2262

Db     181 CTCAGGATGGAGAAGGAGAATCAGAAGCTACAAGCCTCTGAAAACGATTTGCTGAACAAA 240

Qy    2263 cgcctgaagctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgtgg 2322

Db     241 CGCCTCAAGCTTGACTATGATGAAATCACTCCGTGTCTTAAAGAAGTCACTACAGTGTGG 300

Qy    2323 gaaaagatgcttagcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcac 2382

Db     301 GAAAAGATGCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAAGTGCAC 360

Qy    2383 tcggctgttgggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagct 2442

Db     361 TCAGCTGTTGGGCAAGGTGTGCCACGTCATCACCGAGGTGAGATCTGGAAATTTCTAGCT 420

Qy    2443 gagcaattccaccttaaacaccagtttcccagcaaacagcagccaaaggatgtgccatac 2502

Db     421 GAGCAGTTCCACCTTAAACACCCATTTCCTAGTAAACAGCAGCCAAAGGACGTGCCCTAC 480

Qy    2503 aaagaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcga 2562

Db     481 AAAGAGCTCCTGAAGAAGCTGACCTCGCAGCAGCACGCCATTCTCATCGACCTCGGGCGA 540

Qy    2563 acctttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttac 2622

Db     541 ACCTTTCCAACACATCCATACTTCTCTGCCCAGCTTGGAGCAGGTCAGCTGTCACTTTAC 600

Qy    2623 aacattttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagc 2682

Db     601 AACATTCTGAAGGCCTACTCGCTTCTGGACCAGGAGGTTGGATACTGCCAAGGTCTCAGC 660

Qy    2683 tttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaag 2742

Db     661 TTTGTGGCAGGCATTTTGCTTCTTCACATGAGTGAGGAAGAGGCGTTCAAGATGCTCAAG 720

Qy    2743 tttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacag 2802

Db     721 TTCCTGATGTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTATTTTGCAG 780

Qy    2803 atccagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcac 2862

Db     781 ATCCAGATGTACCAGCTGTCACGGCTCCTCCACGATTACCACCGAGACCTCTACAACCAC 840

Qy    2863 ctggaggagcacgagatcggcccagcctctacgctgccccctggttcctcaccatgttt 2922

Db     841 CTGGAAGAGCACGAGATCGGCCCCAGCCTCTACGCGGCTCCCTGGTTTCTCACCGTGTTC 900

Qy    2923 gcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcaggga 2982

Db     901 GCCTCACAGTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCCTTCAGGGA 960

Qy    2983 acagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctg 3042

Db     961 TCAGAGGTCATATTTAAAGTAGCTTTAAGTCTTTTGGGGAGCCATAAGCCCTTGATTCTA 1020

Qy    3043 cagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggc 3102

Db    1021 CAGCATGAGAACCTGGAAACCATCGTGGACTTCATAAAGAACACACTCCCCAACCTGGGC 1080

Qy    3103 ttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagtta 3162

Db    1081 CTGGTGCAGATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTC 1140

Qy    3163 caagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctc 3222

Db    1141 CAGCCCTATGAGGTCGAGTACCACGTGCTCCAGGAGGAGCTTATTGAGTCCTCGCCTCTC 1200

Qy    3223 agtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaac 3282

Db    1201 AGTGACAACCAAAGAATGGAGAAATTGGAGAAAACCAACAGCAGCTTGCGCAAACAGAAC 1260

Qy    3283 cttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccacc 3342

Db    1261 CTTGACCTCCTGGAGCAGTTGCAGGTGGCAAATGCTAGGATCCAAAGCCTTGAAGCCACG 1320

Qy    3343 attgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactg 3402

Db    1321 GTAGAGAAACTTCTTACCAGCGAGAGTAAGCTGAAGCAGGCTGCGCTGACCCTGGAGGTG 1380

Qy    3403 gagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcagagcccagc 3462

Db    1381 GAGCGCTCCGCCCTGCTGCAGATGGTGGAGGAGCTGCGGAGGCAAAGCGCCCGGCCCAGC 1440

Qy    3463 gaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagct 3512

Db    1441 ACTCCAGAGCCAGACTGCACCCAGCTGGAGCCCACAGGCGATTGACCGCT 1490
```

RESULT 8
AK057182
LOCUS AK057182 1788 bp mRNA linear PRI 31-OCT-2001
DEFINITION Homo sapiens cDNA FLJ32620 fis, clone STOMA2000386, highly similar to Bos taurus mRNA for lyncein.
ACCESSION AK057182
VERSION AK057182.1 GI:16552778
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens stomach cDNA to mRNA, clone_lib:STOMA2 clone:STOMA2000386.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1788)
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES Location/Qualifiers
source 1. .1788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STOMA2000386"
/tissue_type="stomach"
/clone_lib="STOMA2"
/note="cloning vector: pME18SFL3"

```
BASE COUNT      481 a    412 c    434 g    461 t
ORIGIN

  Query Match           28.1%;  Score 998.4;  DB 9;  Length 1788;
  Best Local Similarity  99.9%;  Pred. No. 1.3e-207;
  Matches  999;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2556 tgggcgaacctttcctacacacccatacttctctgcccagcttggagcaggacagctatc 2615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    553 TGGGCGAACCTTTCCTACACACCCATACTTCTCTGCCCAGCTTGGAGCAGGACAGCTATC 612

Qy   2616 gctttacaacattttgaaggcctactcacttctagaccaggaagtgggatattgccaagg 2675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    613 GCTTTACAACATTTTGAAGGCCTACTCACTTCTAGACCAGGAAGTGGGATATTGCCAAGG 672

Qy   2676 tctcagctttgtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaat 2735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    673 TCTCAGCTTTGTAGCAGGCATTTTGCTTCTTCATATGAGTGAGGAAGAGGCGTTTAAAAT 732

Qy   2736 gctcaagtttctgatgtttgacatggggctgcggaaacagtatcggccagacatgattat 2795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    733 GCTCAAGTTTCTGATGTTTGACATGGGGCTGCGGAAACAGTATCGGCCAGACATGATTAT 792

Qy   2796 tttacagatccagatgtaccagctctcgaggttgcttcatgattaccacagagacctcta 2855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    793 TTTACAGATCCAGATGTACCAGCTCTCGAGGTTGCTTCATGATTACCACAGAGACCTCTA 852

Qy   2856 caatcacctggaggagcacgagatcggccccagcctctacgctgccccctggttcctcac 2915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    853 CAATCACCTGGAGGAGCACGAGATCGGCCCCAGCCTCTACGCTGCCCCCTGGTTCCTCAC 912

Qy   2916 catgtttgcctcacagttcccgctgggattcgtagccagagtctttgatatgatttttct 2975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    913 CATGTTTGCCTCACAGTTCCCGCTGGGATTCGTAGCCAGAGTCTTTGATATGATTTTTCT 972

Qy   2976 tcagggaacagaggtcatatttaaagtggctttaagtctgttgggaagccataagcccctt 3035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    973 TCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTCTGTTGGGAAGCCATAAGCCCTT 1032

Qy   3036 gattctgcagcatgaaaacctagaaaccatagttgactttataaaaagcacgctacccaa 3095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1033 GATTCTGCAGCATGAAAACCTAGAAACCATAGTTGACTTTATAAAAAGCACGCTACCCAA 1092

Qy   3096 ccttggcttggtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaa 3155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1093 CCTTGGCTTGGTACAGATGGAAAAGACCATCAATCAGGTATTTGAAATGGACATCGCTAA 1152

Qy   3156 acagttacaagcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctc 3215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1153 ACAGTTACAAGCTTATGAAGTTGAGTACCACGTCCTTCAAGAAGAACTTATCGATTCCTC 1212

Qy   3216 tcctctcagtgacaaccaaagaatggataaattagagaaaaccaacagcagcttacgcaa 3275
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1213 TCCTCTCAGTGACAACCAAAGAATGGATAAATTAGAGAAGACCAACAGCAGCTTACGCAA 1272

Qy   3276 acagaaccttgacctccttgaacagttgcaggtggcaaatggtaggatccaaagccttga 3335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1273 ACAGAACCTTGACCTCCTTGAACAGTTGCAGGTGGCAAATGGTAGGATCCAAAGCCTTGA 1332

Qy   3336 ggccaccattgagaagctcctgagcagtgagagcaagctgaagcaggccatgcttacctt 3395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1333 GGCCACCATTGAGAAGCTCCTGAGCAGTGAGAGCAAGCTGAAGCAGGCCATGCTTACCTT 1392

Qy   3396 agaactggagcggtcggccctgctgcagacggtggaggagctgcggcggcggagcgcaga 3455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1393 AGAACTGGAGCGGTCGGCCCTGCTGCAGACGGTGGAGGAGCTGCGGCGGCGGAGCGCAGA 1452

Qy   3456 gcccagcgaccgggagcctgagtgcacgcagcccgagcccacgggcgactgacagctctg 3515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1453 GCCCAGCGACCGGGAGCCTGAGTGCACGCAGCCCGAGCCCACGGGCGACTGACAGCTCTG 1512

Qy   3516 caggagagattgcaacaccatcccacactgtccaggcctt 3555
          |||||||||||||||||||||||||||||||||||||||
Db   1513 CAGGAGAGATTGCAACACCATCCCACACTGTCCAGGCCTT 1552


RESULT   9
AB011175
LOCUS       AB011175                5922 bp    mRNA    linear   PRI 10-APR-1998
DEFINITION  Homo sapiens mRNA for KIAA0603 protein, complete cds.
ACCESSION   AB011175
VERSION     AB011175.1  GI:3043729
KEYWORDS    KIAA0603 protein.
SOURCE      Homo sapiens male brain cDNA to mRNA, clone_lib:pBluescriptII SK
            plus clone:HG1488b.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5922)
  AUTHORS   Ohara,O., Nagase,T. and Ishikawa,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
            DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
REFERENCE   2  (sites)
  AUTHORS   Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
            Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes. IX.
            The complete sequences of 100 new cDNA clones from brain which can
            code for large proteins in vitro
  JOURNAL   DNA Res. 5 (1), 31-39 (1998)
  MEDLINE   98290545
FEATURES             Location/Qualifiers
     source          1. .5922
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HG1488b"
                     /sex="male"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            348. .4247
                     /gene="KIAA0603"
     CDS             348. .4247
                     /gene="KIAA0603"
                     /codon_start=1
                     /product="KIAA0603 protein"
                     /protein_id="BAA25529.1"
                     /db_xref="GI:3043730"
                     /translation="MEPPSCIQDEPFPHPLEPEPGVSAQPGPGKPSDKRFRLWYVGGS
                     CLDHRTTLPMLPWLMAEIRRRSQKPEAGGCGAPAARE VILVLSAPFLRCVPAPGAGAS
                     GGTSPSATQPNPAVFIFEHKAQHISRFIHNSHDLTYFAYLIKAQPDDPESQMACHVFR
                     ATDPSQVPDVISSIRQLSKAAMKEDAKPSKDNEDAFYNSQKFEVLYCGKVTVTHKKAP
                     SSLIDDCMEKFSLHEQQRLKIQGEQRGPDPGEDLADLEVVVPGSPGDCLPEEADGTDT
                     HLGLPAGASQPALTSSRVCFPERILEDSGFDEQQEFRSRCSSVTGVQRRVHEGSQKSQ
                     PRRRHASAPSHVQPSDSEKNRTMLFQVGRFEINLISPDTKSVVLEKNFKDISSCSQGI
                     KHVDHFGFICRESPEPGLSQYICYVFQCASESLVDEVMLTLKQAFSTAAALQSAKTQI
                     KLCEACPMHSLHKLCERIEGLYPPRAKLVIQRHLSSLTDNEQADIFERVQKMKPVSDQ
                     EENELVILHLRQLCEAKQKTHVHIGEGPSTISNSTIPENATSSGRFKLDILKNKAKRS
                     LTSSLENIFSRGANRMRGRLGSVDSFERSNSLASEKDYSPGDSPPGTPPASPPSSAWQ
                     TFPEEDSDSPQFRRRAHTFSHPPSSTKRKLNLQDGRAQGVRSPLLRQSSSEQCSNLSS
                     VRRMYKESNSSSSLPSLHTSFSAPSFTAPSFLKSFYQNSGRLSPQYENEIRQDTASES
                     SDGEGRKRTSSTCSNESLSVGGTSVTPRRISWRQRIFLRVASPMNKSPSAMQQQDGLD
                     RNELLPLSPLSPTMEEEPLVIFLSGEDDPEKIEERKKSKELRSLWRKAIHQQILLLRM
                     EKENQKLEGASRDELQSRKVKLDYEEVGACQKEVLITWDKKLLNCRAKIRCDMEDIHT
                     LLKEGVPKSRRGEIWQFLALQYRLRHRLPNKQQPPDISYKELLKQLTAQQHAILVDLG
                     RTFPTHPYFSVQLGPGQLSLFNLLKAYSLLDKEVGYCQGISFVAGVLLLHMSEEQAFE
                     MLKFLMYDLGFRKQYRPDMMSLQIQMYQLSRLLHDYHRDLYNHLEENEISPSLYAAPW
                     FLTLFASQFSLGFVARVFDIIFLQGTEVIFKVALSLLSSQETLIMECESFENIVEFLK
                     NTLPDMNTSEMEKIITQVFEMDISKQLHAYEVEYHVLQDELQESSYSCEDSETLEKLE
                     RANSQLKRQNMDLLEKLQVAHTKIQALESNLENLLTRETKMKSLIRTLEQEKMAYQKT
                     VEQLRKLLPADALANCDLLLRDLNCNPNNKAKIGNKP"
BASE COUNT     1624 a   1408 c   1432 g   1458 t
ORIGIN
```

Query Match 19.2%; Score 682.2; DB 9; Length 5922;
Best Local Similarity 56.6%; Pred. No. 1.9e-138;
Matches 1585; Conservative 0; Mismatches 1028; Indels 186; Gaps 10;





```
Db    3456 TATGACCTCGGCTTCCGCAAGCAGTACAGACCTGACATGATGTCGCTGCAGATTCAAATG 3515

Qy    2812 taccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctggaggag 2871

Db    3516 TACCAGCTGTCCAGGCTCCTTCATGACTATCACAGAGATCTCTACAATCACCTTGAAGAA 3575

Qy    2872 cacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcctcacag 2931

Db    3576 AATGAAATCAGCCCCAGTCTTTATGCTGCCCCCTGGTTCCTCACATTGTTTGCCTCTCAG 3635

Qy    2932 ttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaacagaggtc 2991

Db    3636 TTTTCATTAGGATTTGTAGCCAGAGTTTTTGATATTATTTTTCTTCAGGGAACTGAAGTT 3695

Qy    2992 atatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcagcatgaa 3051

Db    3696 ATATTCAAGGTTGCACTCAGCCTACTGAGCAGCCAAGAGACACTTATAATGGAATGTGAG 3755

Qy    3052 aacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttggtacag 3111

Db    3756 AGCTTTGAAAATATTGTTGAGTTTCTTAAAAACACGCTACCTGATATGAATACCTCTGAA 3815

Qy    3112 atggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaagcttat 3171

Db    3816 ATGGAAAAAATTATTACCCAGGTTTTTGAGATGGATATTTCTAAGCAGTTGCATGCCTAT 3875

Qy    3172 gaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagt---gac 3228

Db    3876 GAGGTGGAATATCATGTGCTACAGGATGAGCTTCAGGAATCTTCATATTCCTGTGAGGAT 3935

Qy    3229 aaccaaagaatggataaattagagaaaaccaacagcagcttacgcaaacagaaccttgac 3288

Db    3936 AGTGAAACTTTGGAGAAGCTGGAGAGGGCCAATAGCCAACTGAAAAGACAAAACATGGAC 3995

Qy    3289 ctccttgaacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgag 3348

Db    3996 CTCCTAGAAAAATTACAGGTAGCTCATACTAAAATCCAGGCCTTGGAATCAAACCTGGAA 4055

Qy    3349 aagctcctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcgg 3408

Db    4056 AATCTTTTGACGAGAGAGACCAAAATGAAGTCTTTAATCCGGACCCTGGAACAAGAAAAA 4115

Qy    3409 tcggccctgctgcagacggtggaggagctgcggcggcgg 3447

Db    4116 ATGGCTTATCAAAAGACAGTGGAGCAACTCCGGAAGCTG 4154


RESULT 10
AC009595/c
ID   AC009595   standard; DNA; HTG; 189811 BP.
XX
AC   AC009595;
XX
SV   AC009595.3
XX
DT   30-AUG-1999 (Rel. 60, Created)
DT   03-JUL-2000 (Rel. 64, Last updated, Version 4)
XX
DE   Homo sapiens chromosome 4 clone RP11-392K14 map 4, WORKING DRAFT SEQUENCE,
DE   28 unordered pieces.
XX
KW   HTG; HTGS_DRAFT; HTGS_PHASE1.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-189811
RA   Birren B., Linton L., Nusbaum C., Lander E.;
RT   "Homo sapiens chromosome 4, clone RP11-392K14";
RL   Unpublished.
XX
RN   [2]
RP   1-189811
RA   Birren B., Linton L., Nusbaum C., Lander E., Allen N., Anderson M.,
RA   Baker J., Baldwin J., Barna N., Beckerly R., Benn J., Brown A., Castle A.,
RA   Cerny J., Colangelo M., Collins S., Collymore A., Cooke P., DeArellano K.,
RA   Depayre E., Devon K., Dewar K., Donelan L., Doyle M., Ferreira P.,
RA   FitzHugh W., Forrest C., Funke R., Gage D., Galagan J., Gardyna S.,
RA   Gilbert D., Grant G., Hagos B., Heaford A., Horton L., Howland J.C.,
RA   Jones C., Kann L., Karatas A., Lehoczky J., Lieu C., Locke K.,
RA   Macdonald P., Marquis N., McEwan P., McGurk A., McKernan K., McLaughlin J.,
RA   Meldrim J., Molla M., Morris W., Morrow J., Mychaleckyj J., Naylor J.,
RA   Niloff M., O'Connor T., O'Donnell P., Pavlin B., Peterson K., Pollara V.,
RA   Riley R., Roberts D., Roy A., Severy P., Stange-Thomann N., Stojanovic N.,
RA   Stone C., Subramanian A., Tesfaye S., Torruella-Miller I., Vassiliev H.,
RA   Vo A., Wagner A., Wheeler J., Wu X., Wyman D., Ye W.J., Zody M.;
RT   ;
RL   Submitted (27-AUG-1999) to the EMBL/GenBank/DDBJ databases.
RL   Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,
RL   Cambridge, MA 02141, USA
XX
CC   On Jun 23, 2000 this sequence version replaced gi:6006228.
CC   All repeats were identified using RepeatMasker:
CC   Smit, A.F.A. & Green, P. (1996-1997)
CC   http://ftp.genome.washington.edu/RM/RepeatMasker.html
CC   -------------- Genome Center
CC   Center: Whitehead Institute/ MIT Center for Genome Research
CC   Center code: WIBR
CC   Web site: http://www-seq.wi.mit.edu
CC   Contact: sequence_submissions@genome.wi.mit.edu
CC   -------------- Project Information
CC   Center project name: L1755
CC   Center clone name: 392_K_14
CC   -------------- Summary Statistics
CC   Sequencing vector: M13; M77815; 100% of reads
CC   Chemistry: Dye-primer-amersham; 5% of reads
CC   Chemistry: Dye-terminator Big Dye; 95% of reads
CC   Assembly program: Phrap; version 0.960731
CC   Consensus quality: 164704 bases at least Q40
CC   Consensus quality: 175594 bases at least Q30
CC   Consensus quality: 181024 bases at least Q20
CC   Insert size: 214000; agarose-fp
CC   Insert size: 187111; sum-of-contigs
CC   Quality coverage: 3.0 in Q20 bases; agarose-fp
CC   Quality coverage: 3.5 in Q20 bas.
CC   * NOTE: This is a 'working draft' sequence. It currently
CC   * consists of 28 contigs. The true order of the pieces
CC   * is not known and their order in this sequence record is
CC   * arbitrary. Gaps between the contigs are represented as
CC   * runs of N, but the exact sizes of the gaps are unknown.
CC   * This record will be updated with the finished sequence
CC   * as soon as it is available and the accession number will
CC   * be preserved.
CC   *        1      451: contig of 451 bp in length
CC   *      452 551: gap of      100 bp
CC   *      552     2663: contig of 2112 bp in length
CC   *     2664 2763: gap of      100 bp
CC   *     2764     4124: contig of 1361 bp in length
CC   *     4125 4224: gap of      100 bp
CC   *     4225     6767: contig of 2543 bp in length
CC   *     6768 6867: gap of      100 bp
CC   *     6868     8750: contig of 1883 bp in length
CC   *     8751 8850: gap of      100 bp
CC   *     8851    11272: contig of 2422 bp in length
CC   *    11273 11372: gap of      100 bp
CC   *    11373    13411: contig of 2039 bp in length
CC   *    13412 13511: gap of      100 bp
CC   *    13512    16962: contig of 3451 bp in length
CC   *    16963 17062: gap of      100 bp
CC   *    17063    22544: contig of 5482 bp in length
CC   *    22545 22644: gap of      100 bp
CC   *    22645    26284: contig of 3640 bp in length
CC   *    26285 26384: gap of      100 bp
```

```
CC   *    26385    31223: contig of 4839 bp in length
CC   *    31224 31323: gap of    100 bp
CC   *    31324    36014: contig of 4691 bp in length
CC   *    36015 36114: gap of    100 bp
CC   *    36115    41939: contig of 5825 bp in length
CC   *    41940 42039: gap of    100 bp
CC   *    42040    48122: contig of 6083 bp in length
CC   *    48123 48222: gap of    100 bp
CC   *    48223    52983: contig of 4761 bp in length
CC   *    52984 53083: gap of    100 bp
CC   *    53084    58094: contig of 5011 bp in length
CC   *    58095 58194: gap of    100 bp
CC   *    58195    64606: contig of 6412 bp in length
CC   *    64607 64706: gap of    100 bp
CC   *    64707    71079: contig of 6373 bp in length
CC   *    71080 71179: gap of    100 bp
CC   *    71180    79886: contig of 8707 bp in length
CC   *    79887 79986: gap of    100 bp
CC   *    79987    86155: contig of 6169 bp in length
CC   *    86156 86255: gap of    100 bp
CC   *    86256    95370: contig of 9115 bp in length
CC   *    95371 95470: gap of    100 bp
CC   *    95471   103712: contig of 8242 bp in length
CC   *   103713 103812: gap of    100 bp
CC   *   103813   113565: contig of 9753 bp in length
CC   *   113566 113665: gap of    100 bp
CC   *   113666   123867: contig of 10202 bp in length
CC   *   123868 123967: gap of    100 bp
CC   *   123968   132296: contig of 8329 bp in length
CC   *   132297 132396: gap of    100 bp
CC   *   132397   141110: contig of 8714 bp in length
CC   *   141111 141210: gap of    100 bp
CC   *   141211   161622: contig of 20412 bp in length
CC   *   161623 161722: gap of    100 bp
CC   *   161723   189811: contig of 28089 bp in length.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .189811
FT                   /chromosome="4"
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /map="4"
FT                   /clone="RP11-392K14"
FT                   /clone_lib="RPCI-11 Human Male BAC"
FT   misc_feature    1. .451
FT                   /note="assembly_fragment clone_end:T7 vector_side:right"
FT   misc_feature    552. .2663
FT                   /note="assembly_fragment"
FT   misc_feature    2764. .4124
FT                   /note="assembly_fragment"
FT   misc_feature    4225. .6767
FT                   /note="assembly_fragment"
FT   misc_feature    6868. .8750
FT                   /note="assembly_fragment"
FT   misc_feature    8851. .11272
FT                   /note="assembly_fragment"
FT   misc_feature    11373. .13411
FT                   /note="assembly_fragment"
FT   misc_feature    13512. .16962
FT                   /note="assembly_fragment"
FT   misc_feature    17063. .22544
FT                   /note="assembly_fragment"
FT   misc_feature    22645. .26284
FT                   /note="assembly_fragment"
FT   misc_feature    26385. .31223
FT                   /note="assembly_fragment"
FT   misc_feature    31324. .36014
FT                   /note="assembly_fragment"
FT   misc_feature    36115. .41939
FT                   /note="assembly_fragment"
FT   misc_feature    42040. .48122
FT                   /note="assembly_fragment"
FT   misc_feature    48223. .52983
FT                   /note="assembly_fragment clone_end:SP6 vector_side:left"
FT   misc_feature    53084. .58094
FT                   /note="assembly_fragment"
FT   misc_feature    58195. .64606
FT                   /note="assembly_fragment"
FT   misc_feature    64707. .71079
FT                   /note="assembly_fragment"
FT   misc_feature    71180. .79886
FT                   /note="assembly_fragment"
FT   misc_feature    79987. .86155
FT                   /note="assembly_fragment"
FT   misc_feature    86256. .95370
FT                   /note="assembly_fragment"
FT   misc_feature    95471. .103712
FT                   /note="assembly_fragment"
FT   misc_feature    103813. .113565
FT                   /note="assembly_fragment"
FT   misc_feature    113666. .123867
FT                   /note="assembly_fragment"
FT   misc_feature    123968. .132296
FT                   /note="assembly_fragment"
FT   misc_feature    132397. .141110
FT                   /note="assembly_fragment"
FT   misc_feature    141211. .161622
FT                   /note="assembly_fragment"
FT   misc_feature    161723. .189811
FT                   /note="assembly_fragment"
XX
SQ   Sequence 189811 BP; 52455 A; 40839 C; 40321 G; 53485 T; 2711 other;


  Query Match             13.1%;  Score 465.8;  DB 30;  Length 189811;
  Best Local Similarity   99.6%;  Pred. No. 5.2e-91;
  Matches  467;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     416 aagtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagc 475
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97313 AGGTGCCTGAGATCATCAGCTCCATCCGTCAGGCGGGGAAGATCGCCCGGCAGGAGGAGC 97254

Qy     476 tgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcg 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97253 TGCACTGCCCGTCCGAGTTCGACGACACGTTTTCCAAGAAGTTCGAGGTGCTCTTCTGCG 97194

Qy     536 gccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgaga 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97193 GCCGCGTGACGGTGGCGCACAAGAAGGCTCCGCCGGCCCTGATCGACGAGTGCATCGAGA 97134

Qy     596 agttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaaccccgccccatg 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97133 AGTTCAATCACGTCAGCGGCAGCCGGGGGTCCGAGAGCCCCCGCCCCAACCCGCCCCATG 97074

Qy     656 ccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagc 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97073 CCGCGCCCACAGGGAGCCAGGAGCCTGTGCGCAGGCCCATGCGCAAGTCCTTCTCCCAGC 97014

Qy     716 ccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcg 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97013 CCGGCCTGCGCTCGCTGGCCTTTAGGAAGGAGCTGCAGGATGGGGGCCTCCGAAGCAGCG 96954

Qy     776 gcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaata 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   96953 GCTTCTTCAGCTCCTTCGAGGAGAGCGACATTGAGAACCACCTCATTAGCGGACACAATA 96894

Qy     836 ttgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgat 884
           |||||||||||||||||||||||||||||||||||||||||||||||| |
Db   96893 TTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGGT 96845


RESULT  11
AC021106
LOCUS       AC021106              195108 bp    DNA     linear   PRI 03-JUL-2001
DEFINITION  Homo sapiens clone RP11-177C12, complete sequence.
```

```
ACCESSION   AC021106
VERSION     AC021106.6  GI:14589687
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 195108)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 195108)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 195108)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 3, 2001 this sequence version replaced gi:14018126.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information -----------------
            Center project name: H_NH0177C12
            ------------------------------------------------------.
FEATURES             Location/Qualifiers
     source          1. .195108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-177C12"
BASE COUNT    53372 a  40112 c  41873 g  59751 t
ORIGIN


  Query Match             13.1%;  Score 465.8;  DB 9;  Length 195108;
  Best Local Similarity   99.6%;  Pred. No. 5.2e-91;
  Matches  467;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    416 aagtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagc 475
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180759 AGGTGCCTGAGATCATCAGCTCCATCCGTCAGGCGGGGAAGATCGCCCGGCAGGAGGAGC 180818

Qy    476 tgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcg 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180819 TGCACTGCCCGTCCGAGTTCGACGACACGTTTTCCAAGAAGTTCGAGGTGCTCTTCTGCG 180878

Qy    536 gccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgaga 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180879 GCCGCGTGACGGTGGCGCACAAGAAGGCTCCGCCGGCCCTGATCGACGAGTGCATCGAGA 180938

Qy    596 agttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatg 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180939 AGTTCAATCACGTCAGCGGCAGCCGGGGGTCCGAGAGCCCCCGCCCCAACCCGCCCCATG 180998

Qy    656 ccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagc 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180999 CCGCGCCCACAGGGAGCCAGGAGCCTGTGCGCAGGCCCATGCGCAAGTCCTTCTCCCAGC 181058

Qy    716 ccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcg 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181059 CCGGCCTGCGCTCGCTGGCCTTTAGGAAGGAGCTGCAGGATGGGGGCCTCCGAAGCAGCG 181118

Qy    776 gcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaata 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181119 GCTTCTTCAGCTCCTTCGAGGAGAGCGACATTGAGAACCACCTCATTAGCGGACACAATA 181178

Qy    836 ttgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgat 884
          |||||||||||||||||||||||||||||||||||||||||||||||| |
Db 181179 TTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGGT 181227


RESULT  12
AC108933/c
LOCUS       AC108933              208318 bp    DNA     linear   HTG 01-FEB-2002
DEFINITION  Homo sapiens chromosome 4 clone RP11-392K14, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AC108933 AC009595
VERSION     AC108933.1  GI:18464316
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 208318)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 208318)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 1, 2002 this sequence version replaced gi:8671945.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ---------------- Project Information -----------------
            Center project name: H_NH0392K14
            Drafting center: WIBR
            ------------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; %
            Sequencing vector: plasmid; %
            Chemistry: Dye-primer ET; % of reads
            Chemistry: Dye-terminator Big Dye; % of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 205652 bases at least Q40
            Consensus quality: 206608 bases at least Q30
            Consensus quality: 207022 bases at least Q20
            Insert size: 218000; agarose-fp
            Insert size: 207918; sum-of-contigs
            Quality coverage: 6.45 in Q20 bases; agarose-fp
            Quality coverage: 6.62 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    25746: contig of 25746 bp in length
            *    25747    25846: gap of unknown length
            *    25847    65131: contig of 39285 bp in length
            *    65132    65231: gap of unknown length
            *    65232   105028: contig of 39797 bp in length
            *   105029   105128: gap of unknown length
            *   105129   159307: contig of 54179 bp in length
            *   159308   159407: gap of unknown length
            *   159408   208318: contig of 48911 bp in length.
FEATURES             Location/Qualifiers
     source          1. .208318
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="4"
                     /clone="RP11-392K14"
     misc_feature    1. .25746
                     /note="assembly_name:Contig16"
     misc_feature    25847. .65131
                     /note="assembly_name:Contig17"
     misc_feature    65232. .105028
                     /note="assembly_name:Contig18"
     misc_feature    105129. .159307
                     /note="assembly_name:Contig19"
     misc_feature    159408. .208318
                     /note="assembly_name:Contig20"
BASE COUNT    59195 a  45373 c  44728 g  58621 t    401 others
ORIGIN


  Query Match             13.1%;  Score 465.8;  DB 2;  Length 208318;
  Best Local Similarity   99.6%;  Pred. No. 5.2e-91;
  Matches  467;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    416 aagtgcctgagatcatcagctccatccgtcaggcggggaagatcgcccggcaggaggagc 475
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129553 AGGTGCCTGAGATCATCAGCTCCATCCGTCAGGCGGGGAAGATCGCCCGGCAGGAGGAGC 129494

Qy    476 tgcactgcccgtccgagttcgacgacacgttttccaagaagttcgaggtgctcttctgcg 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129493 TGCACTGCCCGTCCGAGTTCGACGACACGTTTTCCAAGAAGTTCGAGGTGCTCTTCTGCG 129434

Qy    536 gccgcgtgacggtggcgcacaagaaggctccgccggccctgatcgacgagtgcatcgaga 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129433 GCCGCGTGACGGTGGCGCACAAGAAGGCTCCGCCGGCCCTGATCGACGAGTGCATCGAGA 129374

Qy    596 agttcaatcacgtcagcggcagccgggggtccgagagcccccgccccaacccgccccatg 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129373 AGTTCAATCACGTCAGCGGCAGCCGGGGGTCCGAGAGCCCCCGCCCCAACCCGCCCCATG 129314

Qy    656 ccgcgcccacagggagccaggagcctgtgcgcaggcccatgcgcaagtccttctcccagc 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129313 CCGCGCCCACAGGGAGCCAGGAGCCTGTGCGCAGGCCCATGCGCAAGTCCTTCTCCCAGC 129254

Qy    716 ccggcctgcgctcgctggcctttaggaaggagctgcaggatgggggcctccgaagcagcg 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129253 CCGGCCTGCGCTCGCTGGCCTTTAGGAAGGAGCTGCAGGATGGGGGCCTCCGAAGCAGCG 129194

Qy    776 gcttcttcagctccttcgaggagagcgacattgagaaccacctcattagcggacacaata 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129193 GCTTCTTCAGCTCCTTCGAGGAGAGCGACATTGAGAACCACCTCATTAGCGGACACAATA 129134

Qy    836 ttgtgcagcccacagatatcgaggaaaatcgaactatgctcttcacgat 884
          |||||||||||||||||||||||||||||||||||||||||||||||| |
Db 129133 TTGTGCAGCCCACAGATATCGAGGAAAATCGAACTATGCTCTTCACGGT 129085


RESULT  13
AL390202_04/c
WPCOMMENT
  Sequence split into 10 fragments  LOCUS AL390202 Accession AL390202
     Fragment Name      Begin    End
     AL390202_00            1    110000
     AL390202_01       100001    210000
     AL390202_02       200001    310000
     AL390202_03       300001    410000
     AL390202_04       400001    510000
     AL390202_05       500001    610000
     AL390202_06       600001    710000
     AL390202_07       700001    810000
     AL390202_08       800001    910000
     AL390202_09       900001    988176
Continuation (5 of 10) of AL390202 from base 400001 (AL390202 Homo sapiens chromosome 20


  Query Match             11.9%;  Score 424.2;  DB 2;  Length 110000;
  Best Local Similarity   96.0%;  Pred. No. 6.1e-82;
  Matches  435;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 atggaaccaataacattcacagcaaggaaacatctgcttcctaacgaggtctcggtggat 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10365 ATGGAACCAATAACATTCACAGCAAGGAAACATCTGCTTCCTAACGAGGTCTCGGTGGAT 10306

Qy     61 tttggcctgcagctggtgggctccctgcctgtgcattccctgaccaccatgcccatgctg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10305 TTTGGCCTGCAGCTGGTGGGCTCCCTGCCTGTGCATTCCCTGACCACCATGCCCATGCTG 10246

Qy    121 ccctgggttgtggctgaggtgcgaagactcagcaggcagtccaccagaaaggaacctgta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10245 CCCTGGGTTGTGGCTGAGGTGCGAAGACTCAGCAGGCAGTCCACCAGAAAGGAACCTGTA 10186

Qy    181 accaagcaagtccggctttgcgtttcaccctctggactgagatgtgaacctgagccaggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10185 ACCAAGCAAGTCCGGCTTTGCGTTTCACCCTCTGGACTGAGATGTGAACCTGAGCCAGGG 10126

Qy    241 agaagtcaacagtgggatcccctgatctattccagcatctttgagtgcaagcctcagcgt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10125 AGAAGTCAACAGTGGGATCCCCTGATCTATTCCAGCATCTTTGAGTGCAAGCCTCAGCGT 10066

Qy    301 gttcacaaactgattcacaacagtcatgacccaagttacttttgcttgtctgattaaggaa 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10065 GTTCACAAACTGATTCACAACAGTCATGACCCAAGTTACTTTGCTTGTCTGATTAAGGAA 10006

Qy    361 gacgctgtccaccggcagagtatctgctatgtgttcaaagccgatgatcaaacaaaagtg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10005 GACGCTGTCCACCGGCAGAGTATCTGCTATGTGTTCAAAGCCGATGATCAAACAAAAGTA 9946

Qy    421 cctgagatcatcagctccatccgtcaggcgggg 453
          ||||||    | |   |    | ||| | ||
Db   9945 AGTGAGATGGAGATCCAAAAGACTAAGGTGTGG 9913


RESULT  14
DMU50542
LOCUS       DMU50542                3214 bp    mRNA    linear   INV 12-JUN-1996
DEFINITION  Drosophila melanogaster pollux (plx) mRNA, complete cds.
ACCESSION   U50542
VERSION     U50542.1  GI:1373162
KEYWORDS    .
SOURCE      fruit fly strain=Oregon R.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 3214)
  AUTHORS   Zhang,S.D., Kassis,J., Olde,B., Mellerick,D.M. and Odenwald,W.F.
  TITLE     Pollux, a novel Drosophila adhesion molecule, belongs to a family
            of proteins expressed in plants, yeast, nematodes, and man
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3214)
  AUTHORS   Zhang,S.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-1996) Shang-Ding Zhang, Neurogenetics, LNC,
            NINDS, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES             Location/Qualifiers
     source          1. .3214
                     /organism="Drosophila melanogaster"
                     /strain="Oregon R"
                     /db_xref="taxon:7227"
                     /chromosome="3"
                     /map="83c"
                     /tissue_type="trachea; CNS"
     gene            1. .3214
                     /gene="plx"
     CDS             130. .2328
                     /gene="plx"
                     /function="adhesion molecule"
                     /codon_start=1
                     /product="pollux"
                     /protein_id="AAB02200.1"
```

```
                     /db_xref="GI:1373163"
                     /translation="MRKPAKRGKRDAAELRELWRTAIRQTIMLNRMETENAMLQARQN
                     ENELKRIKLDYEEIVPCDKQLIERWEQIIERNSTQIGNKKDPKVLGHAIRTGVPRSKR
                     GDVWTFLAEQHSMNTAPVDTKRFPNFNTPYHMLLKHLTEHQHAIFIDLGRTFPNHQFY
                     KDPLGLGQLSLFNLLKAYSILDPELGYCQGLGFICGVLLLHCDEANSFQLLKHLMFRR
                     NMRTKYLPDMKKFQLQLYQLSRLVKDHLPDLYVWLDQNDVSPTLYAAPWILTVFSSQF
                     PLGFVARVFDLLLFLESSDVIFKFAIALLSVHKQQLLAKDNFEEIMDYLKTVVPKMEHT
                     CMEQIMKLVFSMDIGKQLAEYNVEYNVLQEEITTTNHHLEMLNREKTQNQHLEQQLQF
                     AQSSIAQLETTRSSQQAQITTLQSQVQSLELTIQTLGRYVGQLVEHNPDLELPNEVRR
                     MLQQLDDLDRQRRKPIFTERKIGKSVSVSNSHLGFPLKVLEELTERDELGSPQKQKKEK
                     TPFFEQLRQQQQQHRLNGGGQSSNVGESGSPTPPSRPNRLLDNASARTVMQVKLDELK
                     LPEHVDKFVANIKSPLEVDSGVGTPLSPPSTASNSSGGSIFSRMGYRTTPPALSPLAQ
                     RQSYGVAITTAPCPQHMEEVAPATTMAVMPQEDVEEPQPMHPLSMVGGDVNVRFKGTT
                     QLKSIRPVHHMRAIPLGGVQHPSSTEPAVRVAPVPVELAPPAATATTGRS"
BASE COUNT      862 a    798 c    774 g    780 t
ORIGIN
```

Query Match 8.6%; Score 304.6; DB 3; Length 3214;
Best Local Similarity 56.0%; Pred. No. 6.1e-56;
Matches 621; Conservative 0; Mismatches 479; Indels 9; Gaps 2;

```
Qy   2155 gagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggag 2214
          |||||  | || |||||||  ||  || |||   ||    ||| |||||    | ||||||
Db    169 GAGCTGAGGGAACTGTGGCGCACTGCCATTAGGCAGACTATAATGCTGAACCGCATGGAG 228

Qy   2215 aaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctc 2274
          | ||| ||    | | | ||  | |||||      |   |||||| | || ||
Db    229 ACGGAGAACGCCATGTTGCAGGCACGGCAGAATGAGAACGAGCTAAAGCGCATTAAACTG 288

Qy   2275 gattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgctt 2334
          || || || || |||   |||||    ||  |  |||        |||||  ||||  ||
Db    289 GACTACGAGGAGATTGTACCCTGCGACAAGCAGCTAATCGAGCGATGGGAGCAGATCATT 348

Qy   2335 ---agcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgtt 2391
             |||   |   |    ||  |  |||   ||     ||    ||  |   ||  ||
Db    349 GAGCGCAACTCCACACAGATAGGCAACAAGAAGGATCCCAAAGTCCTGGGCCACGCCATT 408

Qy   2392 gggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattc 2451
          |    || || || ||    | || || ||  ||||||  || || |||||||||
Db    409 CGCACCGGAGTACCGCGATCGAAGCGCGGCGATGTCTGGACCTTCCTGGCTGAGCAGCAT 468

Qy   2452 caccttaa------acaccagtttcccagcaaacagcagccaaaggatgtgccatacaaa 2505
            | | ||      ||    ||   | |    |   |          ||   || ||  |
Db    469 TCCATGAACACGGCACCGGTGGACACAAAGCGATTCCCCAACTTCAATACACCGTATCAC 528

Qy   2506 gaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacc 2565
            ||  | || || || |||  || ||||||||||| | ||||| || ||  | ||
Db    529 ATGCTGCTGAAACACCTTACTGAGCATCAGCATGCGATTTTCATTGATCTGGGCAGGACT 588

Qy   2566 tttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaac 2625
          |||||||  ||||  | || |   |  | ||||||    || ||| | ||||| |  |||
Db    589 TTTCCTAATCACCAGTTCTACAAGGATCCGCTTGGTCTCGGCCAGTTGTCGCTGTTTAAC 648

Qy   2626 attttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagcttt 2685
            |  ||||||| |||||  |||||||  | |||| | || || |||||  |||||  ||||
Db    649 CTGCTGAAGGCGTACTCCATTCTAGATCCGGAATTAGGGTACTGCCAGGGTCTGGGCTTC 708

Qy   2686 gtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagttt 2745
            |   ||| | || || || |||     ||| |   |  | ||| || |||| |||
Db    709 ATCTGCGGCGTCTTACTCCTGCATTGCGATGAAGCCAATTCATTTCAACTGCTGAAGCAC 768

Qy   2746 ctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatc 2805
          || ||||||  |       |||| |    | || | ||| |||||||  |  || ||  |
Db    769 CTAATGTTTCGCCGTAATATGCGCACGAAATACCTGCCGGACATGAAAAAGTTTCAACTG 828

Qy   2806 cagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctg 2865
          ||| |||| ||||||||  | ||| |  | ||| | |    ||| ||||||      ||
Db    829 CAGCTGTATCAGCTCTCCCGATTGGTCAAGGATCATCTGCCAGATCTCTACGTGTGGCTC 888

Qy   2866 gaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcc 2925
          ||  || ||||  ||   ||||   | || || || || ||| |||||||| | |||  |
Db    889 GATCAGAACGATGTCTCGCCCACTTTGTATGCGGCTCCATGGATCCTCACCGTCTTTAGC 948

Qy   2926 tcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaaca 2985
          |||||||| |||||||||||||| ||| | ||||||||| || | || ||  |     |
Db    949 TCACAGTTTCCGCTGGGATTCGTGGCCCGCGTCTTTGATCTGCTCTTCCTGGAATCCTCC 1008

Qy   2986 gaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcag 3045
          || || || |||||  | ||  |      || |        || ||||      | |||
Db    1009 GATGTAATCTTTAAGTTTGCCATTGCCTTGCTCTCCGTGCACAAGCAGCAACTACTGGCC 1068

Qy   3046 catgaaaacctagaaaccatagttgactttataaaaagcacgctacccaaccttggcttg 3105
            | || ||| | |||   ||  | |||| | | || | |  | | || ||  | |
Db    1069 AAGGATAACTTCGAAGAGATTATGGACTATCTGAAGACCGTGGTGCCAAAGATGGAGCAC 1128

Qy   3106 gtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaa 3165
                |||||  | |  || || | ||| |||   |||||||| |  || ||| |
Db    1129 ACGTGCATGGAGCAAATTATGAAGCTGGTCTTTAGCATGGACATTGGAAAGCAGCTTGCC 1188

Qy   3166 gcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagt 3225
          |  ||   | || |||||| | || ||||| || ||  |||  |     |    | || |
Db    1189 GAATACAATGTGGAGTACAATGTGCTTCAGGAGGAGATTACCACCACTAACCATCACCTA 1248

Qy   3226 gacaaccaaagaatggataaattagagaa 3254
          || |  |  |    ||| ||     ||||
Db    1249 GAAATGCTTAACCGGGAGAAGACGCAGAA 1277
```

```
RESULT  15
DME17919
LOCUS       DME17919                4746 bp    mRNA    linear   INV 18-NOV-1998
DEFINITION  Drosophila melanogaster mRNA for pollux protein.
ACCESSION   Y17919
VERSION     Y17919.1  GI:3893102
KEYWORDS    pollux gene.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 4746)
  AUTHORS   Xu,X.Z., Wes,P.D., Chen,H., Li,H.S., Yu,M., Morgan,S., Liu,Y. and
            Montell,C.
  TITLE     Retinal targets for calmodulin include proteins implicated in
            synaptic transmission
  JOURNAL   J. Biol. Chem. 273 (47), 31297-31307 (1998)
  MEDLINE   99030403
REFERENCE   2  (bases 1 to 4746)
  AUTHORS   Montell,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-1998) C. Montell, The Johns Hopkins University
            School of Medicine, Department of Biological Chermistry, 725 N.
            Wolfe Street, Baltimore, MD 21205-2185, USA
FEATURES             Location/Qualifiers
     source          1. .4746
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
     source          1. .3316
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /map="83C"
                     /tissue_type="retinal"
                     /dev_stage="adult"
     gene            452. .4591
                     /gene="pollux"
     CDS             452. .4591
                     /gene="pollux"
                     /note="calmodulin binding protein"
                     /codon_start=1
                     /product="Pollux protein"
                     /protein_id="CAA76939.1"
                     /db_xref="GI:3893103"
                     /translation="MAELMHQMRDPAHTLGGSVGSIPQTLIGGGGGSHGNSNGALNGI
                     HATPATNLKMSEAMRNAQHDTSANPVSSKMKASKSYTHGLSSSSGTVNIPTSTSAQSN
                     LSLLADISPNHTHFFEVMYVGKIRVSQKRVPNTFIDDALPKFKAYDAQRLRLLQNRKM
```

SLSSEGGVGIEAKPSSSLKSHDLKEEDEEEQEQHKGHDDSQDSQAKPLVQLQLTGAEE
GAAPRPLEDNKENKSPEKRPLLRGQSQIELGHKEHSDGSQPSAANSQLEAPNVIVNKQ
PTPPRDQGVGTGTASASAGPSQLHPNYAMDNIPKQRDRSASQGCIPPYVEQNRTMVFL
VGRCDLRLISPDRKQVLLYKDFKDVASCVHGQKSLDHFGIICRELNNDGYIGYVFKCQ
SEHVCDDIVAAIAQAFDTCAEQKKKQDTQIFSCEHCPMLWYHKLCTDVEGLSEKKTQA
LILRRIETLSDDEQEIVWAKFCGSEKTNSPVAEQNQFLMMLLRAHCESRQQRHVHDTA
ENRSEFLNQYLGGSTIFMKAKRSLTNSFDNLLKRKPSKDDIAVPSHNLRDIREGSAEP
LGTQSPPEGFRSRSNTVGASPSSKPTAEQLKSPMMDIFIKVGNSPKEAETHQGSWRQA
ILNSVVTPSKGLDSEVPTEFLSPMRKPAKRGKRDAAELRELWRTAIRQTIMLNRMETE
NAMLQARQNENELKRIKLDYEEIVPCDKQLIERWEQIIERNSTQIGNKKDPKVLGHAI
RTGVPRSKRGDVWTFLAEQHSMNTAPVDTKRFPNFNTPYHMLLKHLTEHQHAIFIDLG
RTFPNHQFYKDPLGLGQLSLFNLLKAYSILDPELGYCQGLGFICGVLLLHCDEANSFQ
LLKHLMFRRNMRTKYLPDMKKFQLQLYQLSRLVKDHLPDLYVWLDQNDVSPTLYAAPW
ILTVFSSQFPLGFVARVFDLLFLESSDVIFKFAIALLSVHKQQLLAKDNFEEIMDYLK
TVVPKMEHTCMEQIMKLVFSMDIGKQLAEYNVEYNVLQEEITTTNHHLEMLNRERTQN
QHLEQQLQFAQSSIAQLETTRSSQQAQITTLQSQVQSLELTIQTLGRYVGQLVEHNPD
LELPNEVRRMLQQLDDLDRQRRKPIFTERKIGKSVSVNSHLGFPLKVLEELTERDELG
SPQKQKKEKTPFFEQLRQQQQQHRLNGGGQSSNVGESGSPTPPSRPNRLLDNASARTV
MQVKLDELKLPEHVDKFVANIKSPLEVDSGVGTPLSPPSTASNSSGGSIFSRMGYRTT
PPALSPLAQRQSYGVAITTAPCPHHMEEVAPATTMAVMPQEDVEEPQPMHPLSMVGGD
VNVRFKGTTQLKSIRPVHHMRAIPLGGVQHPSSTEPAVRVAPVPVELAPPAATATTGR
S"

BASE COUNT 1211 a 1342 c 1261 g 932 t
ORIGIN

Query Match 8.6%; Score 304.6; DB 3; Length 4746;
Best Local Similarity 56.0%; Pred. No. 6.4e-56;
Matches 621; Conservative 0; Mismatches 479; Indels 9; Gaps 2;

```
Qy   2155 gagctccgagagctgtggcaaaaggctattcttcaacagatactgctgcttagaatggag 2214
          |||||  | || |||||||  |  || |||   ||    ||| |||||    | ||||||
Db   2432 GAGCTGAGGGAACTGTGGCGCACTGCCATTAGGCAGACTATAATGCTGAACCGCATGGAG 2491

Qy   2215 aaggaaaatcagaagctccaagcctctgaaaatgatttgctgaacaagcgcctgaagctc 2274
          | ||| ||    | | | || ||     | |||||      |   |||||| | || ||
Db   2492 ACGGAGAACGCCATGTTGCAGGCACGGCAGAATGAGAACGAGCTAAAGCGCATTAAACTG 2551

Qy   2275 gattatgaagaaattactccctgtcttaaagaagtaactacagtgtgggaaaagatgctt 2334
          || || || || |||   |||||    ||  |  |||        |||||  ||||  ||
Db   2552 GACTACGAGGAGATTGTACCCTGCGACAAGCAGCTAATCGAGCGATGGGAGCAGATCATT 2611

Qy   2335 ---agcactccaggaagatcaaaaattaagtttgacatggaaaaaatgcactcggctgtt 2391
             |||   |   |     |  | |||   ||     ||    ||  |   ||  ||
Db   2612 GAGCGCAACTCCACACAGATAGGCAACAAGAAGGATCCCAAAGTCCTGGGCCACGCCATT 2671

Qy   2392 gggcaaggtgtgccacgtcatcaccgaggtgaaatctggaaatttctagctgagcaattc 2451
           |   || || || ||     | || || ||  ||||||  || || ||||||||
Db   2672 CGCACCGGAGTACCGCGATCGAAGCGCGGCGATGTCTGGACCTTCCTGGCTGAGCAGCAT 2731

Qy   2452 caccttaa------acaccagtttcccagcaaacagcagccaaaggatgtgccatacaaa 2505
            | | ||      ||    |  | |    |   |          ||   || ||  |
Db   2732 TCCATGAACACGGCACCGGTGGACACAAAGCGATTCCCCAACTTCAATACACCGTATCAC 2791

Qy   2506 gaactcttaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacc 2565
             ||  | || || || |||  || ||||||||||||| | ||||| || ||  | ||
Db   2792 ATGCTGCTGAAACACCTTACTGAGCATCAGCATGCGATTTTCATTGATCTGGGCAGGACT 2851

Qy   2566 tttcctacacacccatacttctctgcccagcttggagcaggacagctatcgctttacaac 2625
          |||||||  ||||  | || |  |  | ||||||    || ||| | ||||| |  |||
Db   2852 TTTCCTAATCACCAGTTCTACAAGGATCCGCTTGGTCTCGGCCAGTTGTCGCTGTTTAAC 2911

Qy   2626 attttgaaggcctactcacttctagaccaggaagtgggatattgccaaggtctcagcttt 2685
           |  ||||||| ||||  |||||| | |||| | || || ||||| |||||  ||||
Db   2912 CTGCTGAAGGCGTACTCCATTCTAGATCCGGAATTAGGGTACTGCCAGGGTCTGGGCTTC 2971

Qy   2686 gtagcaggcattttgcttcttcatatgagtgaggaagaggcgtttaaaatgctcaagttt 2745
           |    ||| | || || || |||     ||| |   |  | ||| || |||| |||
Db   2972 ATCTGCGGCGTCTTACTCCTGCATTGCGATGAAGCCAATTCATTTCAACTGCTGAAGCAC 3031

Qy   2746 ctgatgtttgacatggggctgcggaaacagtatcggccagacatgattattttacagatc 2805
          || ||||||  |       |||| |   | || | ||| |||||||  |  || ||  |
Db   3032 CTAATGTTTCGCCGTAATATGCGCACGAAATACCTGCCGGACATGAAAAAGTTTCAACTG 3091

Qy   2806 cagatgtaccagctctcgaggttgcttcatgattaccacagagacctctacaatcacctg 2865
          ||| |||| |||||||||  | ||| |  | ||| | |    ||| ||||||       ||
Db   3092 CAGCTGTATCAGCTCTCCCGATTGGTCAAGGATCATCTGCCAGATCTCTACGTGTGGCTC 3151

Qy   2866 gaggagcacgagatcggccccagcctctacgctgccccctggttcctcaccatgtttgcc 2925
          ||  || |||| ||    |||| | || || || || ||| |||||||| | |||  |
Db   3152 GATCAGAACGATGTCTCGCCCACTTTGTATGCGGCTCCATGGATCCTCACCGTCTTTAGC 3211

Qy   2926 tcacagttcccgctgggattcgtagccagagtctttgatatgatttttcttcagggaaca 2985
          |||||||| |||||||||||||| ||| | |||||||| || | || ||  |     |
Db   3212 TCACAGTTTCCGCTGGGATTCGTGGCCCGCGTCTTTGATCTGCTCTTCCTGGAATCCTCC 3271

Qy   2986 gaggtcatatttaaagtggctttaagtctgttgggaagccataagcccttgattctgcag 3045
          || || || |||||  | ||  |     || |       || ||||       | |||
Db   3272 GATGTAATCTTTAAGTTTGCCATTGCCTTGCTCTCCGTGCACAAGCAGCAACTACTGGCC 3331

Qy   3046 catgaaaacctagaaaccatagttgacttttataaaaagcacgctacccaaccttggcttg 3105
           | || ||| | |||    ||  | |||| | | || | |  | | || ||  | |
Db   3332 AAGGATAACTTCGAAGAGATTATGGACTATCTGAAGACCGTGGTGCCAAAGATGGAGCAC 3391

Qy   3106 gtacagatggaaaagaccatcaatcaggtatttgaaatggacatcgctaaacagttacaa 3165
               |||||  | |  || || | ||| |||   |||||||| |  || ||| |
Db   3392 ACGTGCATGGAGCAAATTATGAAGCTGGTCTTTAGCATGGACATTGGAAAGCAGCTTGCC 3451

Qy   3166 gcttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctcctctcagt 3225
          |  ||  | || |||||| | || ||||| || ||  ||| |    |    | || |
Db   3452 GAATACAATGTGGAGTACAATGTGCTTCAGGAGGAGATTACCACCACTAACCATCACCTA 3511

Qy   3226 gacaaccaaagaatggataaattagagaa 3254
          || |  |  |    ||| ||     ||||
Db   3512 GAAATGCTTAACCGGGAGAAGACGCAGAA 3540
```

Search completed: September 9, 2002, 19:16:17
Job time: 17490 sec